

A Decision Support Model for Cloud Bursting

Cloud bursting combines the advantages of private and public clouds by adding external resources when internal resources are insufficient. Private clouds are still less expensive and remain under the control of the owner, but public cloud resources of any size are available on demand. The article suggests a decision support model to compute the optimal amount of internal resources and cost reductions.

DOI 10.1007/s12599-013-0257-5

The Author

Dr. Markus Lilienthal (✉)

E-FinanceLab
RuW-Gebäude, Raum 1.218
Goethe-Universität Frankfurt
Grüneburgplatz 1
60323 Frankfurt am Main
Germany
MarkusLilienthal@gmx.de

Received: 2011-11-28

Accepted: 2012-11-06

Accepted after two revisions by
Prof. Dr. Bichler.

Published online: 2013-02-28

This article is also available in German in print and via <http://www.wirtschaftsinformatik.de>: Lilienthal M (2013) Ein Entscheidungsmodell für Cloud-Bursting. WIRTSCHAFTSINFORMATIK. doi: 10.1007/s11576-013-0357-3.

© Springer Fachmedien Wiesbaden 2013

1 Introduction

Cloud computing (Weinhardt et al. 2009) has gained notable popularity in the past few years. With this paradigm, users may buy computing resources as a scalable utility service. Gartner Market Research estimates that this infrastructure, as a service market, will grow from \$3.7 billion in 2011 to \$10.5 billion in 2014 (Petty 2011). Already in 2011, cloud computing has become a top technology priority (Luftman and Zadeh 2011; Moore 2011). Yet many enterprises still have doubts about public cloud computing (Petty and Tudor 2010) and prefer to rely on private clouds, constrained to one enter-

prise or to a cluster of trustworthy enterprises (i.e., community clouds). Cloud bursting (Kailasam et al. 2010) is sometimes used synonymously with the terms augmented cloud computing (Assunção et al. 2009; Reynolds and Bess 2009) or hybrid clouds (Armbrust et al. 2010; Mateescu et al. 2011). It offers a further, mixed strategy that can offload some workload onto public clouds (external providers) and thus obtain infrastructure as a service when internal resources are not sufficient.

Mainstream adoption of cloud bursting is likely five to ten years away (Smith 2011); the adoption of cloud computing remains a complex decision. This article focuses on the economic aspect of cloud computing in order to develop an economic decision model that determines the optimal size of internal resources and the resulting cost savings achieved through cloud bursting. Consideration of workload distributions and especially demand volatility are central to this issue. The model relies on a modification of classical lease-or-buy models to address this requirement. The model recognizes that public cloud resources usually are more expensive than internal resources when fully utilized. Therefore, bursting is the economically optimal strategy, on condition that the right size of the private (internal) cloud resources is chosen.

As a purely economic model, the model presumes that cloud bursting is managerially feasible (outsourcing decision) and that there are suitable providers that can offer required service levels (vendor selection problem). Several studies in information systems research already address these questions as separate decision problems, also with specific focus on cloud computing.

As a practical demonstration, the empirical section examines two applications. The first dataset serves to determine the optimal size of a datacenter

needed by a large bank if it uses cloud bursting. The bank's low current utilization rate allows the cloud bursting introduction to generate cost savings of approximately 70 %. The second dataset shows the potential created by pooling three workloads under cloud bursting, in order to compare the effects of pooling and bursting. The three unpooled clouds, when subjected to cloud bursting, would be approximately 50 % more expensive than the current pooled datacenter without bursting. But pooled workloads with cloud bursting are 6% less expensive.

2 Related Literature

2.1 Cloud Computing

The U.S. National Institute of Standards and Technology (Mell and Grance 2009) defines cloud computing as a pay-per-use model that provides available, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services), which can be rapidly provisioned and released with minimal management effort or service provider interaction. This definition comprises five key characteristics: on-demand self-service, ubiquitous network access, location-independent resource pooling, rapid elasticity, and pay-per-use. Vaquero et al. (2009) also cite resource virtualization, scalability through seemingly infinite resources, customized service level agreements, ease of use, and commercial pay-per-use models as commonly accepted key characteristics. These characteristics appear as presumptions in the proposed model. The model however relaxes the pay-per-use requirement, which is more a business model than a technical characteristic. A descriptive review of cloud computing literature is

available in Yang and Tate (2009). The focus of this paper is more an economic than a technical one. Therefore, the next sections of related literature concentrate on papers with this orientation.

2.2 Lease-or-Buy Decisions

From an economic point of view, the decision to use cloud bursting relates to a lease-or-buy problem. Lease-or-buy problems are frequent subjects of investigation in management and accounting research, and operations literature. Almost everything, from car pools to copiers, can be either leased or bought.

The key difference between classical lease-or buy problems and the model described in this article is that they typically do not consider any mixed strategies. Also, the volatile demand as key driver is not in focus. Early solutions appear in Weekes et al. (1969), Harwood and Hermanson (1976), and Anderson and Martin (1977).

Moving on to contemporary, cloud-related publications, Walker et al. (2010) develop a decision model applied to storage cloud applications and based on the results of Johnson and Lewellen (1972). Their approach calculates the separate net present values (NPV) of the leasing decision and the buying decision. The NPV sums all current and future costs or earnings while discounting the future cash flows, which results in future cash flows counting less than present cash flows. Thus, the NPV models a preference for regular costs over immediate one-off costs. The discount is modeled like a negative interest rate, such that constant regular cash flows always have a finite NPV.

There are a few other publications that have analyzed the costs of cloud computing as a classical lease-or-buy problem: Deelman et al. (2008) analyze a specific data-intensive application, whereas Walker (2009) compares the lifetime costs of leasing cloud resources to the costs of purchasing resources, expressed as the cost per CPU hour. Walker concludes that leasing is the less expensive option if the lifetime of purchased resources is extremely long (10 years or more) or utilization is low (40 % scenario). Finally, Li et al. (2009) develop a cost calculation tool to determine the total cost of cloud ownership which can be used to determine the input for such a decision model. However, none of these authors considers a mixed strategy (cloud bursting) or stochastic demand.

The classical lease-or-buy decision does not consider all problem dimensions. This paper will also not discuss every problem dimension. For instance, it does not consider additional managerial effects of outsourcing. Extensive research summaries about outsourcing in general and its research directions can be found in Dibbern et al. (2004) or Lacity and Willcocks (2000). More recent research focuses on cloud computing as a specific case of outsourcing (Blaskovich and Mintchik 2011; Lacity et al. 2010; Motahari-Nezhad et al. 2009; Yang 2011).

Also, the classical lease-or-buy typically does not consider the vendor selection. Finding an optimal vendor is not trivial and may influence the lease-or-buy decision itself. Vendor selection is a traditional research area in operations research, decision science, marketing, and management, which generally consider manufactured goods. In an often-cited milestone, Dickson (1966) identifies criteria used by purchasing managers. Several recent contributions apply and concretize results from classical vendor selection literature in the context of cloud computing (Garg et al. 2011; Godse and Mulik 2009; Kang and Sim 2010; Li et al. 2011; Rehman et al. 2011).

2.3 Cloud Bursting

Cloud bursting as a special form of cloud computing is a central concept in this paper. This last literature section clarifies its role and definition. The term cloud bursting was originally used to describe an extension of grids and clusters by means of clouds, and thus a way of transition toward cloud computing (Marshall et al. 2010; Mateescu et al. 2011). Reynolds and Bess (2009) even classify cloud bursting as an adoption strategy, which they call the replacement strategy. Yet cloud bursting also offers an alternative to leasing versus buying not only during the adoption process. Cloud bursting facilitates the shift of infrastructure-as-a-service workload to public clouds (external providers) when internal resources are insufficient (Goyal 2010; Kailasam et al. 2010; Raj 2011). The question of pooling several private clouds owned by trusted entities in one community cloud may coincide with the decision to adopt cloud bursting or offer an alternative to it. Technical approaches for implementing cloud bursting comprise the OpenNebula project (Cerbelaud et al. 2009) and Aneka (Vecchiola et al. 2011).

There are only a handful of empirical studies and decision models for cloud migration. Several case studies have reported on the practical experiences of migrating to the cloud. For example, Khajeh-Hosseini et al. (2010) calculate cost savings and other stakeholder benefits, such as satisfaction, for two servers and find cost savings of 37 %. Assunção et al. (2009) simulate the cost/benefits of cloud bursting as a cluster extension by determining an optimal scheduling strategy. The economic decision model for cloud bursting of software applications that Strebel and Stage (2010) develop focuses on optimal resource allocations (internal/external) as a mixed-integer programming problem. The model demands a fixed number of hosted applications as input and does not consider stochastic demand.

Bibi et al. (2010) instead collect relevant drivers of cloud adoption and propose a three-step decision model (assessment of software and infrastructure development costs, definition of quality characteristics, and estimation of user demand). However, they do not detail how the three steps actually lead to a decision. They also structure the various sources of costs and how to estimate them. Their classification of costs as fixed or variable provides a valuable guide for performing the preliminary steps of the proposed method.

Another elaborate decision tool for migration to the cloud, the Cloud Adoption Toolkit (Khajeh-Hosseini et al. 2012), consists of technology suitability analyses, cost modeling, and stakeholder impact analyses. The first element is a simple questionnaire that checks necessary preconditions; cost modeling is based on deterministic workload patterns. Thus, the Cloud Adoption Toolkit can model time-dependent demand but not random demand peaks. Nor does it consider a bursting strategy as a mixture of leasing or buying. Finally, consulting firms have addressed the cloud migration problem, too; Forrester suggest bursting as a strategic option in their decision framework (Staten 2009a, 2009b). However, such frameworks typically are proprietary and not completely transparent.

Thus, most of the existing contributions propose a general framework that maps the complete adoption decision, but does not detail the economic perspective or merely analyzes a strict lease-or-buy decision for cloud computing without including the option of bursting.

Some approaches disentangle the components of fixed and variable costs, but only one research paper truly addresses the economic aspect of cloud bursting. It does not consider a stochastic workload distribution, but rather focuses on mixed-integer programming to obtain an optimal allocation for a fixed number of concurrent applications.

3 Model Description

The model presented here attempts to minimize the expected costs that arise as a sum of four cost components. This expected value approach therefore considers anticipated costs per interval. The model differentiates internal capacity (private resources) and external resources offered by a commercial provider. The size of the internal capacity is subject to optimization, whereas external resources are on demand. The model allows for extreme cases of no internal capacity or no external resources if we set the respective variables to 0. The concept of cloud bursting means that external resources can be bought whenever internal capacity is insufficient. Furthermore, the proposed model can be extended or altered to fit additional requirements or relax assumptions. For example, it might be helpful to consider cancellation costs instead of buying external resources (Beck et al. 2008) or to turn to multidimensional distributions that relax the assumption of one resource unit.

3.1 Assumptions

The model relies on several assumptions and preconditions. First, cloud bursting should be already assessed as technically and managerially feasible. Khajeh-Hosseini et al. (2012) show how such considerations can be structured. Second, the model assumes only one resource metric (e.g., number of standard virtual instances). In typical cases, cloud resources sold as infrastructure-as-a-service use standard instances as unit, and the user must determine the required number of instances, depending on the most critical underlying resource. It is the task of automated schedulers to deploy these instances as physical resources. In the model described in this article, internal and external resources differ only in price. That is, providers without acceptable service levels do not enter the consideration set.

Table 1 Model variables

<i>Decision variable</i>	
Cap	Internal capacity in workload units
<i>Input variables</i>	
F	Cumulative workload distribution function (workload units per interval)
$c_{in,fix}$	Internal fixed costs (per capacity unit)
$c_{in,var}$	Internal variable costs (per workload unit)
$c_{ex,fix}$	External fixed costs
$c_{ex,var}$	External variable costs (per workload unit)
All	Allowance (workload units per interval)

Third, by definition, the supply of external resources are not limited in cloud computing. Because users thus do not perceive shortfalls, demand does not depend on supply. This means that users do not consume more resources just because a large amount of resources are currently unused, or vice versa. A more technical assumption is that resources are continuously scaled (not just integer values). This assumption is an acceptable simplification for applications with sufficient demand. If not applicable, continuous distributions can be replaced by discrete distributions.

With regard to practical application, application resource scaling can be done in general either horizontally or vertically. Horizontal scaling means applying more resources of the same kind in parallel. Vertical scaling means applying more powerful resources. Cloud computing as infrastructure service is primarily a means of horizontal scaling. Thus, in order to harvest the bursting capabilities it is necessary to have applications or demand bursts that can be horizontally scaled (i.e., can be distributed). With the diffusion of the cloud technology increasingly more applications offer the functionality of horizontal scaling.

3.2 Workload Distribution

Workload is the amount of necessary resources per time interval, e.g. number of instances per hour. A reasonable length for the time intervals is often the typical minimum lease time for external resources. For example, Amazon EC2, a popular provider of cloud services (Amazon 2012), currently allows the creation of additional instances within minutes. However, the minimum lease time and billing increment is one hour. Based on this knowledge, one hour represents an appropriate interval. The resources will

be measured in *resource units* (workload units or capacity units), which could be standard instances as a very abstract unit or a typical critical unit (depending on the application) such as number of CPUs or storage units.

Workload is a random variable, i.e., it has a stochastic distribution. The *workload distribution* is the distribution of the amount of resources needed within one interval. Formally, we denote the workload distribution function as F (Table 1).

The number of potentially appropriate distributions is vast (Feitelson 2005), and the best choice depends on the application. Heavy-tailed distributions should be reasonably truncated to avoid infinite expected values. A feasible approach is to use standard distribution functions (log-normal, Gamma, Weibull, and Pareto) to avoid over-fitting. Workload distributions vary from case to case, depending on the organizational environment, scheduling and priority handling, usage and user characteristics, system architecture, and time of day/week/year (Feitelson 2005). Therefore, the distribution needs to be determined on the basis of data. There are different options for obtaining such a workload distribution. For example, the IT department could simply ask departments (users) how many resources they typically require and how much that amount typically varies. Because metering in the proposed model is very high-leveled, users likely can cope with such estimation. With the method of moments (Lublin and Feitelson 2003), only a few parameters need to be estimated, given a certain distribution class. A more time-consuming but also more accurate method to obtain workload distributions would be to use scheduler traces or less elaborate utilization logging tools.

The number of available private cloud resources will be called *internal capacity*.

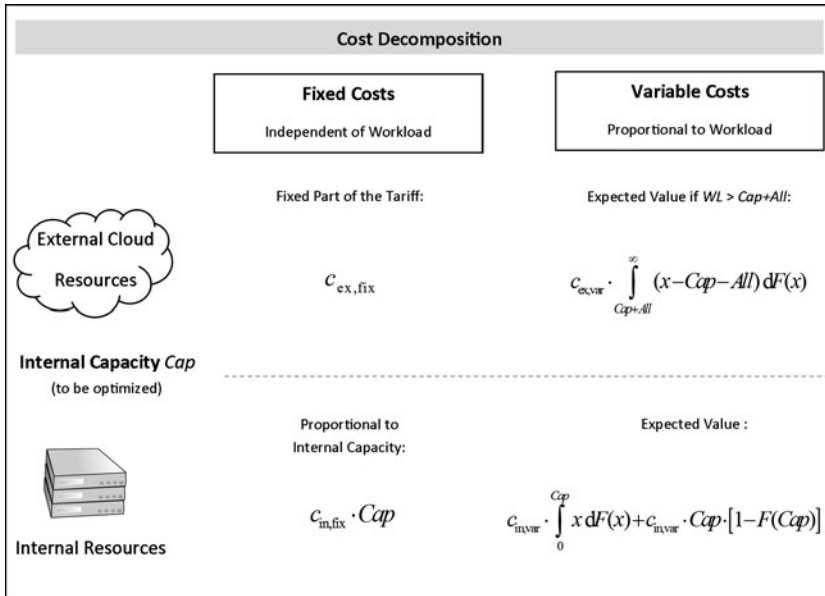


Fig. 1 Four cost components (see also **Table 1**)

Internal capacity is a fixed value and the decision variable in the model.

3.3 Costs

The model differentiates four types of costs per interval (**Table 1**), indicated with the subscript of the cost parameter c . By transitioning from costs per unit to expected costs per interval, given the workload distribution F , the total expected costs are the sum of the four components displayed in **Fig. 1**. The optimization task is to minimize the sum of these four expected costs with respect to internal capacity Cap as the decision variable.

In the first dimension (lower vs. upper part of **Fig. 1**), costs split into internal (i.e., private cloud solution) and external (i.e., those occurring when internal resources do not suffice) costs. Using internal resources to a certain extent is cost-optimal only if the internal costs under high utilization conditions are lower than the external costs. This is typically true (Opitz et al. 2008; Walker 2009; Walker et al. 2010). In the second dimension, costs split into variable and fixed costs (left vs. right). Variable costs depend on the amount actually consumed; fixed costs do not depend on the actual workload.

The internal fixed costs (lower left, **Fig. 1**) depend on the size of the internal capacity, not on actual utilization. These costs include the one-off costs of investments divided by their lifetime, as well

as utilization-independent running costs, such as the costs of computing hardware and software, business premises, personnel, and communication. Most of these costs are proportional to the internal capacity, within a certain scale. Therefore, internal variable costs are proportional to internal capacity in the model.

The internal variable costs (lower right, **Fig. 1**) consist of all internal costs that are proportional to actual usage. They come into play if, for example, energy-saving measures reduce costs at low utilization rates. The internal variable costs are proportional to the actual internal capacity used, and total internal capacity (Cap) is the upper limit. The summand $c_{in,var} \cdot \int_0^{Cap} x dF(x)$ represents expected costs when the workload is less than internal capacity, and the summand $c_{in,var} \cdot Cap \cdot [1 - F(Cap)]$ represents the expected costs for the rest of the time.

The external cost structure is determined by the tariff. Although many cloud computing definitions include pay-per-use as a key characteristic (Vaquero et al. 2009), market developments and also customer preferences (Koehler et al. 2010) indicate increasing popularity of multipart tariffs, and also researchers predict that prices will become more dynamic and efficient (e.g., Beck et al. 2008). Both IBM and Amazon currently sell instances as either pay-per-use or “reserved instances” that are actually two-part tariffs: After paying a fixed fee (monthly or yearly), the variable costs for

using each instance are lower. The general idea of offering such tariffs, from the provider’s perspective, is to earn higher profits through higher consumption. Often, these tariffs also benefit customers in form of consumer surplus.

The proposed model incorporates three-part tariffs: a fixed component $c_{ex,fix}$ (e.g., monthly fee), a free consumption allowance All , and variable costs $c_{ex,var}$ (cf. Lambrecht et al. 2007). By setting the allowance equal to 0, we obtain the special case of a two-part tariff. If the entrance fee is 0, we obtain a pay-per-use equation. Thus, the model offers the flexibility to incorporate different tariffs by just setting the cost variables appropriately.

4 Optimal Internal Capacity in the Model

The optimal strategy for how much internal capacity Cap should be provided comes from minimizing the total expected costs as a function of internal capacity. The total costs equal the sum of the four components in **Fig. 1**:

$$\begin{aligned}
 Costs(Cap) &= c_{in,fix} \cdot Cap + c_{in,var} \cdot \int_0^{Cap} x dF(x) \\
 &+ c_{in,var} \cdot Cap \cdot [1 - F(Cap)] \\
 &+ c_{ex,fix} + c_{ex,var} \\
 &\cdot \int_{Cap+All}^{\infty} (x - Cap - All) dF(x) \quad (1)
 \end{aligned}$$

The second summand can be rewritten as

$$\begin{aligned}
 c_{in,var} \cdot \int_0^{Cap} x dF(x) &= c_{in,var} \cdot \int_0^{\infty} x dF(x) \\
 &- c_{in,var} \cdot \int_{Cap}^{\infty} x dF(x) \\
 &= c_{in,var} \cdot EV(WL) \\
 &- c_{in,var} \cdot \int_{Cap}^{\infty} x dF(x) \quad (2)
 \end{aligned}$$

where $EV(WL)$ denotes the expected value of the workload. Together,

$$\begin{aligned}
 Costs(Cap) &= c_{in,fix} \cdot Cap + c_{in,var} \cdot EV(WL) \\
 &+ c_{in,var} \cdot Cap \cdot [1 - F(Cap)] \\
 &+ c_{ex,fix} + c_{ex,var} \\
 &\cdot \int_{Cap+All}^{\infty} (x - Cap - All) dF(x)
 \end{aligned}$$

$$- c_{in,var} \cdot \int_{Cap}^{\infty} x dF(x) \quad (3)$$

The roots of the derivative yield locally optimal capacities:

$$\begin{aligned} Costs'(Cap) &= c_{in,fix} - c_{in,var} \cdot Cap \cdot f(Cap) \\ &+ c_{in,var} \cdot [1 - F(Cap)] \\ &- c_{ex,var} [1 - F(Cap + All)] \\ &+ c_{in,var} \cdot Cap \cdot f(Cap) \\ &= c_{in,fix} + c_{in,var} \cdot [1 - F(Cap)] \\ &- c_{ex,var} [1 - F(Cap + All)] \quad (4) \end{aligned}$$

The parameters $c_{in,fix}$ and $c_{ex,var}$ are typically the most relevant and most prominently considered costs; that is, internal capacities are governed by fixed costs, independent of consumption, whereas the external cloud offers are pay per use ($All = 0$ and $c_{ex,fix} = 0$). Disregarding the zero terms ($c_{in,var} = c_{ex,fix} = 0$), optimal internal capacity in this condition is obtained by finding the root of this derivative

$$\begin{aligned} 0 &= c_{in,fix} - c_{ex,var} [1 - F(Cap + All)] \\ &\Leftrightarrow \\ Cap^* &= F^{-1} \left(1 - \frac{c_{in,fix}}{c_{ex,var}} \right) \quad (5) \end{aligned}$$

where F^{-1} is the inverse distribution function. Therefore, the optimal internal capacity depends on the ratio of internal fixed costs and external variable costs, as well as a specific quantile of the workload distribution determined on this ratio. With an increasing ratio, the optimal internal capacity decreases—whereas the role of the distribution is more complex, as explained below.

Consider a hypothetical numerical example. Assume that the workload distribution is simply normal, with a mean of 1,000 units and a standard deviation of 300 units. The costs in the bursting case are defined as $c_{in,fix} = 1$ and $c_{ex,var} = 2$ (i.e., external costs are twice as high as the internal costs; cf. Armbrust et al. (2010) or Shroff (2010)). Then, the optimal internal capacity for cloud bursting is 1,000 units (the median), with a 50 % probability of bursting events (5). The corresponding costs are 1,000 units for the internal resources, plus mean external costs of 239 units (numerical solution of the upper right part of Fig. 1), for a total of 1,239 units. Let us assume the enterprise has a private cloud of 2,116 resource

units at the current time, such that current resources are sufficient with a probability of 99.99 % (four-nine availability). The available resources are used at less than half their capacity on average, and cost 2,116 units. The exclusive use of external resources instead costs 2,000 units on average (= mean workload \times external variable costs). In this stylized numerical example, cloud bursting would save roughly 40 % of the costs evoked through either full leasing or no leasing. In the latter case, it also avoids shortages.

How does expected workload and variance (volatility) influence the optimal capacity in (5)? The optimal internal capacity increases if expected workload increases, *ceteris paribus* (i.e., variance remains the same). Increasing the expected workload means shifting the density function along the x-axis to the right. Thus, all quantiles shift in the same direction, especially optimal capacity as a quantile.

For details about the effect of variance (volatility), *ceteris paribus*, consider a transformed workload $\frac{x-\mu}{\sigma} \tilde{\sigma} + \mu$ instead of x with distribution \tilde{F} . This expression has the same expected value μ but a different (say, higher) standard deviation ($\tilde{\sigma}$ instead of σ). Thus, \tilde{F} has higher variance, *ceteris paribus*. The transformation increases all observations (and all quantiles) beyond the expected value and reduces all observations (and all quantiles) smaller than the expected workload. Because Cap^* is higher than expected workload, switching from F to \tilde{F} increases the optimal internal capacity (and reduces it if the condition is not satisfied). The optimal capacity is higher than the expected workload if the external costs are high enough. The threshold depends on the distribution and its shape. Marketing research notes, in a slightly different context, the presence of similar behavior: Under three-part subscription tariffs, uncertain consumers tend to choose higher allowances (Lambrecht et al. 2007).

When we switch from the consideration of pay-per-use to two-part tariffs ($c_{ex,fix} > 0$), we find that optimal capacity is not affected if the external variable costs $c_{ex,var}$ remain equal. In practice, two-part tariffs have lower variable costs, as an incentive for more consumption. Thus the optimal capacity for two-part tariffs, according to (5), is lower. If we introduce an allowance ($All > 0$), the optimal capacity decreases by this allowance:

$$Cap^* = F^{-1} \left(1 - \frac{c_{in,fix}}{c_{ex,var}} \right) - All. \quad (6)$$

Therefore, the allowance directly increases the amount of consumed external resources. Because this tariff is usually more closely associated with lower variable costs, the optimal capacity is overall lower. If two or more tariffs are available, the actual costs must be compared to determine which the cheaper tariff is. If $c_{ex,fix}$ is too high, then two- and three-part tariffs are not attractive. Finally, we consider variable internal costs $c_{in,var}$. If there is no allowance ($All = 0$), the external variable costs in (5) get replaced by the difference between external and internal variable costs:

$$Cap^* = F^{-1} \left(1 - \frac{c_{in,fix}}{c_{ex,var} - c_{in,var}} \right). \quad (7)$$

If the internal fixed costs were to transform into internal variable costs, optimal internal capacity would increase. The effect of volatility on the optimal internal capacity is more likely negative though, because the ratio in the brackets is larger. The remaining general case in which all parameters are positive has no closed-form solution, but the effects of the parameters offer a mixture of the discussed effects.

5 Applications

5.1 Description

The first empirical application is a pure application of the model (computation of the optimal capacity and cost savings obtained by introducing cloud bursting). The computation follows the numerical example in the previous section but is based on real data. The second empirical application however is an example of how the model can be used to derive additional insights. An enterprise may save money by pooling the (internal) resources of different departments or creating a community cloud with other enterprises (Mohan 2011). This option can be combined with cloud bursting or regarded as a competing strategy. The proposed model can also compute the effects on costs in both cases.

The cost savings then depend on the correlation of the workloads. In the worst case scenario, two departments have a perfect correlation of +1. In the best case, the two departments are negatively correlated (i.e., the workload of one department is high, and the other's is low). Both cases are relevant in practice: Workloads of different departments in the same

time zone are probably positively correlated, whereas follow-the-sun scenarios (Beulen et al. 2005), which aggregate workloads from different departments in different time zones, are negatively correlated.

If the workloads of n departments have distributions F_n , the expected value of the sum of those workloads is equal to the sum of the expected values μ_n (if they exist). The variance is the sum of all pairwise covariances (entries of the covariance matrix). Therefore, negative correlation (negative covariance) leads to lower variance in the aggregated workload.

For some standard distributions, such as normal ones, the means and covariance provide sufficient information to determine the distribution of the total workload easily (i.e., sums of normal distributions are normal). If not, the method of moments can approximate the distribution of the total workload (for log-normal distributions, see Schwartz and Yeh (1982)). Below the applied method is Monte Carlo sampling. The procedure in application 1 also applies to separate workloads, and the total workload yields the cost savings of pooling.

5.2 Data Sets

The first (application 1) is a performance log of a large international bank's cloud datacenter, which contains 96 uniform nodes. The servers are primarily used for risk analysis. The dataset lists hourly CPU utilization rates in October 2008, measured with the RRD tool (Oetiker 2011). In comparison to other logged resources CPU utilization is the most critical resource. The bank also provided monthly aggregated overall cost calculations for six months, which allowed estimating the approximate costs per hour per node, through linear regression. The 95 % confidence interval of the costs per server per hour is bound by €0.27 and €0.39 (regardless of actual utilization rate). The average utilization was approximately 14 %, and the highest recorded utilization was only 67 % (Table 2). The distribution parameters indicate the institution is risk averse in terms of resource shortfalls.

The data in application 2 are trace logs of the Auvergne Grid (Anoep et al. 2011). The measured resource is again CPU utilization. Grids can provide an infrastructure for clouds or be extended by clouds

Table 2 Descriptive statistics, first dataset (percentage utilization)

N	Mean	Std. dev.	Min	5 %	95 %	Max
744	13.5 %	10.9 %	0.8 %	1.4 %	32.6 %	67.1 %

Table 3 Descriptive statistics, second dataset (by group, percentage utilization)

Group	Mean	Std. dev.	Min	5 %	95 %	Max
Altogether	60.5 %	25.5 %	0.0 %	12.1 %	97.4 %	171.1 %
A	16.9 %	14.3 %	0.0 %	0.6 %	45.8 %	76.7 %
B	28.2 %	22.6 %	0.0 %	0.0 %	65.6 %	85.6 %
C	15.4 %	21.8 %	0.0 %	0.0 %	60.8 %	136.7 %

through cloud bursting (Marshall et al. 2010; Mateescu et al. 2011). The scientific production grid has 405 users and a high utilization rate. The analysis concentrates on a subset of data from April to October that is characterized by a high utilization of 61% on average (Table 3, first row). The dataset offers both submission and waiting times of jobs. The analysis considers the submission times and thus eliminates scheduling effects. Because of that, the maximum workload (Table 3, first row) is beyond the 100 % threshold. This dataset provides the necessary information to reveal the effects of pooling and correlations across users or user groups. Again, the chosen time intervals have a length of one hour, which is much less than the average observed waiting time (162 minutes) and in agreement with the minimum lease time of Amazon EC2 (Amazon 2012).

The user group with highest demand accounts for 47 % of the total workload (group B, Table 3). The group with second highest demand (group A, Table 3) accounts for slightly more than half of that amount. The other groups have only small and irregular workloads. They will be summarized altogether as group C. Group C has much higher demand peaks than the other two groups. The workloads of the three groups are significantly and negatively correlated (Table 4). Apparently, there is some additional scheduling effect beyond the technical level. Therefore, one analysis scenario re-estimates the aggregated workload of the different user groups assumed the single group workload distributions are independent. In contrast with the first dataset, the Auvergne Grid does not provide cost data. So, similar to the approach of Strebel and Stage (2010), the model assumes cost ratios that reflect

Table 4 Pairwise correlations of user groups (significant at $p < .001$)

Group	A	B
B	-0.19	
C	-0.11	-0.36

prior literature and the observation of first dataset.

5.3 Execution and Results: Application 1

The software EasyFit (MathWave Technologies 2010) automatically estimates up to 65 different distributions (mostly using maximum likelihood methods) and ranks them according to Kolmogorov-Smirnov, Anderson-Darling, and Chi-squared fit tests. From the results, three very common standard distributions with appropriate domains and moderate fit (Weibull, Gamma, log-normal) have been selected, as well as the very flexible, well-fitting Wakeby distribution, to show the sensitivity of the distribution choice. The Wakeby distribution is an adjustable distribution with five parameters (Rao and Hamed 1999) that typically provides better fit than the other distributions. The other three distributions have only two parameters; the log-normal distribution also has a long tail that emphasizes demand peaks. Across these three fit tests, the Wakeby and Weibull distributions are the best fitting distributions with appropriate domains. The log-normal distribution, with its long tail, suffers the poorest fit, as is visible in Fig. 2; apparently, it overemphasizes demand peaks.

The applied internal costs are the average of €0.33 per hour, independent of utilization. The pay-per-use costs

of external resources depend on the provider. Amazon (2012), the most popular provider, sold standard instances from less than \$1 to more than \$2 per hour at the time of conduction. The computational example compares three different levels of external costs: €1.00, €1.50, and €2.00.

Figure 3 shows the resulting optimal capacity for the three cost levels (determined by (5)) and a graphical illustration of the price sensitivity for the four distributions. The optimal capacity increases with higher external costs. These results do not differ much across the four distributions, which suggests the model is robust to the choice of the distribution. The most differing values are those of the log-normal distribution with the apparently poorest fit.

The optimal capacity appears to be only around 20 % of current capacity – a result of the very low average utilization. We observe that the stronger emphasis on the distribution tail by the log-normal

distribution leads to slightly higher optimal internal capacity if external prices are large but a lower optimal internal capacity if external prices are low.

Figure 4 provides the corresponding cost savings at optimal capacity as a percentage of the costs under current capacity. The cost savings are high because the observed utilization is low; for example, Khajeh-Hosseini et al. (2010) report a cost reduction of only 37 %. The cost savings of approximately 70 % are somewhat lower than the reduction of internal capacities (approximately 80 %), because external resources are more expensive. Cost savings also decrease with higher external costs. More apparent than in Fig. 3, the log-normal distribution differs; the other three distributions together yield robust results.

5.4 Execution and Results: Application 2

The standard condition in this application will be again $c_{ex,var}/c_{in,fix} = 5$, comparable to the result in the first dataset

and those obtained by other authors (e.g., Armbrust et al. 2010). With the second dataset, we can disaggregate users to analyze cost savings through pooling. The analysis compares three hypothetical cases that pool resources differently. First, we consider three disaggregated cases in which all three user groups (A, B, and C; Table 3) have their own datacenters. Second, we analyze the observed aggregated workload (with a slightly negative correlation). Third, we assess a similar hypothetical case but with zero correlation among user groups. To obtain this condition, the independent demand sum of the three distributions in the first case is simulated and the results are fitted again (Monte Carlo solution). The independent sum approximates an elimination of the observed scheduling above the technical level.

The workload distributions of the three groups have different characteristics (Table 5). Group B compared with group A has a flat distribution, with a high probability of no workload. The aggregated distribution compared with the separate groups offers a clear mode (inflection point of the cumulative distribution function). The independent sum reveals a characteristic distribution tail after the elimination of the top-level scheduling. Similar to the first dataset, the EasyFit application (MathWave Technologies 2010) lists the Wakeby distribution as the best fitting distribution in almost all cases (Table 5), though its results are not notably different from those of the Weibull or Gamma distributions. The log-normal distribution, however, again provides obviously the worst fit.

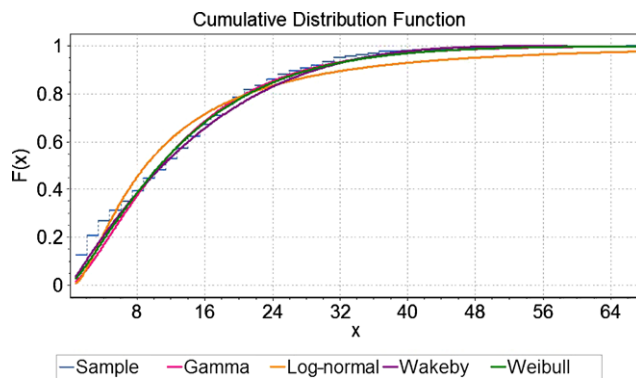


Fig. 2 Observed distribution and four fitted distributions

Optimal capacity	External costs		
	€1.00	€1.50	€2.00
Weibull	16 %	20 %	23 %
Gamma	15 %	20 %	23 %
Wakeby	17 %	21 %	24 %
Log-normal	14 %	20 %	24 %

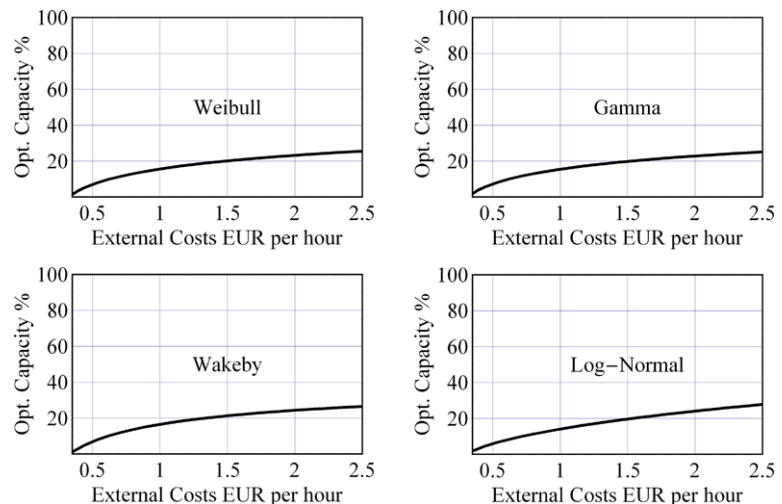


Fig. 3 Optimal capacity (percentage of current) for different levels of external costs and distributions

Optimal savings	External costs		
	€1.00	€1.50	€2.00
Weibull	74 %	70 %	67 %
Gamma	74 %	70 %	67 %
Wakeby	74 %	70 %	68 %
Log-normal	68 %	60 %	54 %

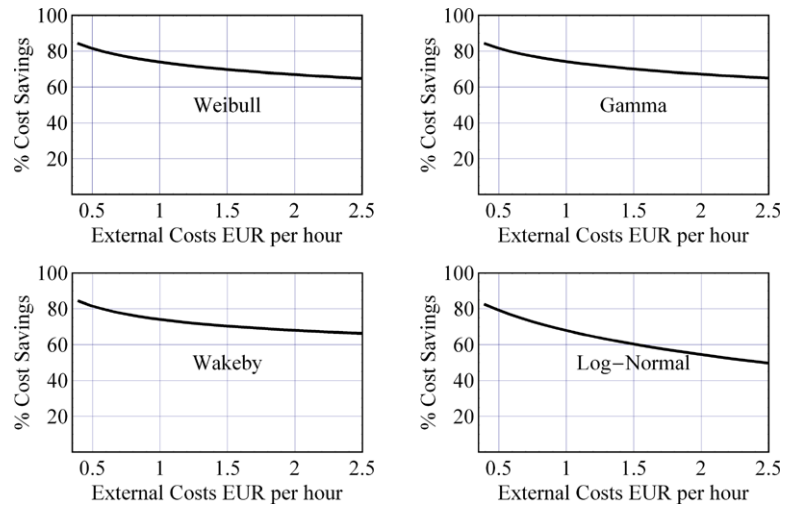


Fig. 4 Cost savings (percentage of current) for different levels of external costs and distributions

Table 5 Distribution fit for application 2

	Group A	Group B	Group C	Altogether	Independent Sum
	2/2/1	5/3/NA	14/1/2	1/1/1	3/38/NA
Wakeby					
Weibull	7/19/10	30/31/16	19/17/32	18/19/18	12/13/12
Gamma	18/28/21	22/23/24	8/14/11	25/25/20	20/17/17
Log-normal	24/32/32	26/35/32	20/30/26	40/37/41	29/30/30

Notes: Group A, B and C refer to the workloads of the three groups separately, the forth column is the observed workload altogether, and the fifth column contains the simulated independent sum. Each cell offers a graphical comparison of observed and fitted distribution functions and the rank among all 65 tested distributions (rank 1 = best fit, Kolmogorov-Smirnov/Anderson-Darling/Chi-square)

Table 6 offers the results of the decision model applying the Wakeby distribution as the best fitting distribution and $c_{ex,var} : c_{in,fix} = 5 : 1$. The internal costs at actual capacity are 100 % = 1. Without

cloud bursting, it would be equal to the actual costs. Optimal capacity with cloud bursting (**Table 6**, fourth row) is 22 % lower than actual (100 %), and cloud bursting would save 6 % of costs com-

pared with the actual costs without cloud bursting. Even with the negative correlation eliminated (last row), the optimal capacity remains below actual capacity. In this case however, external resources

are more intensively consumed, so total costs are higher than actual costs. This effect comes from the additional external resources that are needed to compensate the higher volatility that was partly suppressed by the negative correlation.

The total costs of separate clouds for each group ($40 + 62 + 52 = 154$ %) are even higher. That is, the cost savings obtained from pooling in this dataset equal $(154 - 115)/154 = 25$ % for uncorrelated workloads or $(154 - 94)/154 =$

39 % with the observed negative correlation.

Finally, **Figs. 5 and 6** show the sensitivity of the observed correlated total workload to different distributions and costs (both referring to the fourth row in **Table 6**). If the ratio of external costs and internal costs is not too high, the differences among the Weibull, Gamma, and Wakeby distributions are small. As for data set 1, the optimal capacity (**Fig. 6**) is more robust with respect to the distribution assumption than the total costs. The differences in optimal capacity are small for small and medium external costs, except in the log-normal distribution. Again, the log-normal distribution with its poor fit renders much different results and should not be considered. The sensitivity analyses of the other results

Table 6 Estimated optimal capacity and cost savings by separate groups, with correlated (actual), and with artificially independent workloads. Costs at actual capacity without cloud bursting = 100 %

Scenario	Optimal capacity (percentage of current)	Total costs
Group A only	28 %	40 %
Group B only	50 %	62 %
Group C only	24 %	52 %
All correlated (observed)	78 %	94 %
All independent (assumed)	86 %	115 %

Total costs	External costs		
	3:1	5:1	7:1
Weibull	90 %	98 %	103 %
Gamma	89 %	100 %	106 %
Wakeby	86 %	94 %	99 %
Log-normal	122 %	148 %	167 %

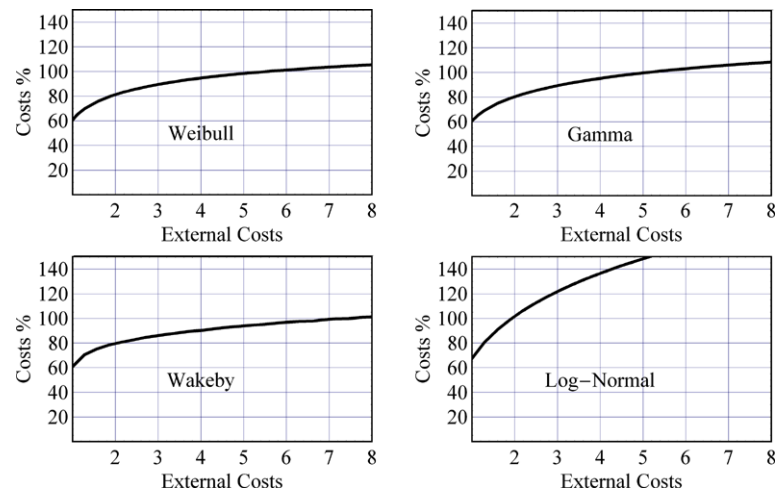


Fig. 5 Total costs under the correlated workload (**Table 6**, fourth row) for different levels of external costs and distributions

Optimal capacity	External costs		
	3:1	5:1	7:1
Weibull	71 %	82 %	89 %
Gamma	68 %	80 %	87 %
Wakeby	70 %	78 %	83 %
Log-normal	71 %	95 %	111 %

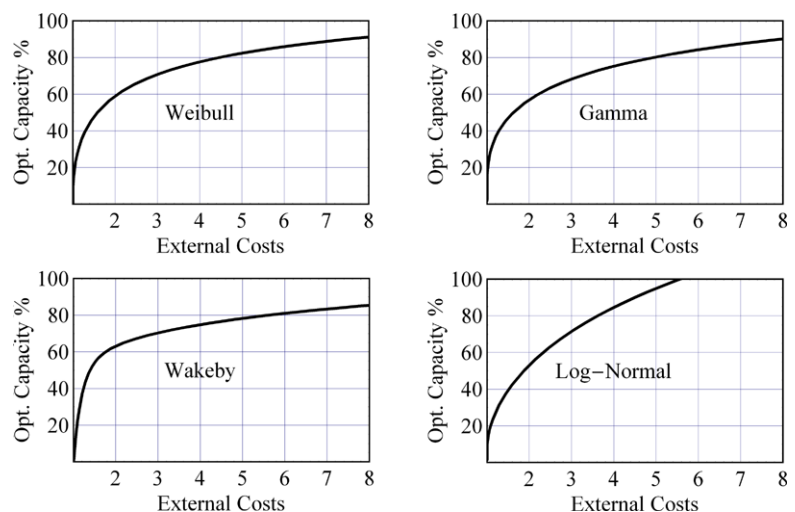


Fig. 6 Optimal capacity under the correlated workload (**Table 6**, fourth row) for different levels of external costs and distributions

Abstract

Markus Lilienthal

A Decision Support Model for Cloud Bursting

The cloud computing market divides into public (commercial) and private (self-provisioned) clouds. The concept of cloud bursting combines public and private clouds: The private cloud (internal resources) provides the computational capacity, but a part of the demand is offloaded onto public clouds. This article proposes an easy-to-apply economic decision support model for determining on the one hand the optimal size of the internal capacity for cloud bursting technology, and on the other hand the cost savings. The model uses an expected value approach that considers stochastic workload and is flexible with respect to the distribution choice. Two empirical examples demonstrate the applicability of the model.

Keywords: Cloud computing, Cloud bursting, Augmented cloud computing, Lease or buy, Computing economics, Cost savings, Cloud adoption

in **Table 6** are not displayed but yield qualitatively the same conclusions.

5.5 Summary of the Results

The first application, a purely empirical elaboration of the model to determine the optimal internal capacity, shows a high cost saving potential of 70 %. This large value coincides with a low observed average utilization of only 14 %. The second application further investigates the role of pooling when applying the model to either separated or aggregated data. The analyzed system is well-utilized at 61 %, thus the cost savings of cloud bursting with respect to the aggregated (and slightly negatively correlated) workload is small (6 %). However, the cost savings from pooling are considerable: Pooling all three user groups saves 39 % of the costs compared with employing individually owned cloud bursting solutions.

The results in both applications also are mostly robust to the choice of the distribution, though the distribution should have at least moderate fit (as is true for three of the four distributions in both applications). Higher external costs lead to a higher optimal internal capacity and lower cost savings, but small errors in costs do not bias the results substantially.

6 Conclusions

The idea of cloud bursting provides an opportunity to combine the appeals of private cloud resources with the benefits of buying additional resources if private resources are not sufficient. This research offers a flexible decision model to describe the economic trade-off between private clouds and public clouds, with cloud bursting as a third option. The decision model is flexible with respect to the workload distribution.

The article presents two empirical applications. The first uses workload data from a large international bank, with separately provided cost data. Such datasets are rare; from that perspective, the results are insightful. The analysis yields predicted cost savings of 70 % by using cloud bursting, compared with the current costs that accrue from the very low level of observed utilization. The second application uses the model to also determine the effects of pooling workloads. The cost savings then depend on the degree of correlation across user pools.

Three separate clouds, each with cloud bursting, would incur 154 % of the current costs and 102 % of current capacity (without bursting). All three groups pooled together, though, require only 94 % of the current costs at 78 % of the current capacity under cloud bursting. The result depends on the assumed cost structure, which was estimated for this application.

The applications represent two specific cases. For example, the existing utilization in application 1 is very low, which leads to high cost savings. It is difficult to generalize these empirical results. Yet the intention of these two empirical studies is to demonstrate the applicability of the model, not provide representative figures. The two applications reveal that the benefits differ strongly from case to case but can potentially be quite significant.

The theoretical section of this research notes the effects of average resource consumption, volatility, and pricing of external resources on an optimal strategy. As long as the price ratio between external and internal resources is large, greater volatility leads to an increase in the amount of internal resources as optimal value. However, in the case of a small ratio, demand volatility becomes a driver of the use of external resources. This observation confirms the widespread belief that public cloud computing is particularly beneficial for small enterprises with volatile resource demand and relatively high costs of running their own hardware.

The model also confirms that the optimum quantity of internal resources diminishes with multipart tariffs. Therefore, cloud providers should follow the example of the telecommunications industry when transitioning from simple pay-per-use to multipart tariffs. Marketing literature has shown that multipart tariffs benefit not only providers but can also benefit customers through greater consumer surplus.

Finally, the value of the proposed model lies in its applicability to complex situations and its flexibility in terms of workload patterns (distributions). Unlike existing models, it considers stochastic demand. Similar studies can easily be conducted by IT departments.

References

Amazon (2012) Amazon EC2 pricing. <http://aws.amazon.com/ec2/pricing/>. Accessed 2012-03-20

- Anderson PF, Martin JD (1977) Lease vs purchase decisions: a survey of current practice. *Financial Management* 6(1):41–47
- Anoop S, Dumitrescu C, Epema D, Iosup A, Jan M, Li H, Wolters L (2011) The grid workloads archive. TU Delft, Delft
- Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, Lee G, Patterson D, Ariel Rabkin IS, Zaharia M (2010) A view of cloud computing. *Commun ACM* 53(4):50–58
- Assunção MD, Costanzo A, Buyya R (2009) Evaluating the cost-benefit of using cloud computing to extend the capacity of clusters. In: International symposium on high performance distributed computing (HPDC), Munich
- Beck R, Schwind M, Hinz O (2008) Grid economics in departmentalized enterprises. *Journal of Grid Computing* 6(3):277–290
- Beulen E, Fenema P, Currie W (2005) From application outsourcing to infrastructure management: extending the offshore outsourcing service portfolio. *European Management Journal* 23(2):133–144
- Bibi S, Katsaros D, Bozani P (2010) Application development: fly to the clouds or stay in-house? In: 19th IEEE international workshops on enabling technologies: infrastructures for collaborative enterprises, Larissa, pp 60–65
- Blaskovich J, Mintchik N (2011) Information technology outsourcing: a taxonomy of prior studies and directions for future research. *Journal of Information Systems* 25(1):1–36
- Cerbelaud D, Garg S, Huylebroeck J (2009) Opening the clouds: qualitative overview of the state-of-the-art open source VM-based cloud management platforms. In: 10th ACM/IFIP/USENIX international conference on middleware, Urbana Campaign
- Deelman E, Singh G, Livny M, Berriman B, Good J (2008) The cost of doing science on the cloud: the montage example. In: 2008 ACM/IEEE conference on supercomputing, Austin
- Dibbern J, Goles T, Hirschheim R, Jayatilaka B (2004) Information systems outsourcing: a survey and analysis of the literature. *The Data Base for Advances in Information Systems* 35(4):6–102
- Dickson G (1966) An analysis of vendor selection systems and decisions. *Journal of Purchasing* 2(1):5–17
- Feitelson DG (2005) Workload modeling for computer systems performance evaluation. Unpublished book draft
- Garg SK, Versteeg S, Buyya R (2011) SMICloud: a framework for comparing and ranking cloud services. In: Fourth IEEE international conference on utility and cloud computing, Victoria, pp 210–218
- Godse M, Mulik S (2009) An approach for selecting software-as-a-service (SaaS) product. In: IEEE international conference on cloud computing, Bangalore
- Goyal P (2010) Enterprise usability of cloud computing environments: issues and challenges. In: 19th IEEE international workshops on enabling technologies: infrastructures for collaborative enterprises, Larissa, pp 54–59
- Harwood GB, Hermanson RH (1976) Lease-or-buy decisions. *Journal of Accountancy* 142(3):83–87
- Johnson RW, Lewellen WG (1972) Analysis of the lease-or-buy decision. *Journal of Finance* 27(4):815–823
- Kailasam S, Gnanasambandam N, Dhara-nipragada J, Sharma N (2010) Optimizing service level agreements for autonomic cloud bursting schedulers. In: 39th international conference on parallel processing, workshops, San Diego
- Kang J, Sim KM (2010) Cloudle: a multi-criteria cloud service search engine. In: IEEE Asia-Pacific services computing conference, Hangzhou
- Khajeh-Hosseini A, Greenwood D, Sommerville I (2010) Cloud migration: a case study of migrating an enterprise IT system to IaaS. In: 3rd international conference on cloud computing, Miami
- Khajeh-Hosseini A, Greenwood D, Smith JW, Sommerville I (2012) The cloud adoption toolkit: supporting cloud adoption decisions in the enterprise. *Software: Practice and Experience* 42(4):447–465
- Koehler P, Anandasivam A, Dan MA, Weinhardt C (2010) Customer heterogeneity and tariff biases in cloud computing. In: 31st international conference on information systems, Saint Louis
- Lacity MC, Willcocks L (2000) Global information technology outsourcing. In: Search of business advantage. Wiley, New York
- Lacity MC, Khan S, Yan A, Willcocks LP (2010) A review of the it outsourcing empirical literature and future research directions. *Journal of Information Technology* 25(4):395–433
- Lambrecht A, Seim K, Skiera B (2007) Does uncertainty matter? Consumer behavior under three-part tariffs. *Marketing Science* 26(5):698–710
- Li A, Yang X, Kandula S, Zhang M (2011) Comparing public-cloud providers. *IEEE Internet Computing* 15(2):50–53
- Li X, Li Y, Liu T, Qiu J, Wang F (2009) The method and tool of cost analysis for cloud computing. In: 2nd IEEE international conference on cloud computing, Bangalore
- Lublin U, Feitelson DG (2003) The workload on parallel supercomputers: modeling the characteristics of rigid jobs. *Journal of Parallel and Distributed Computing* 63(11):1105–1122
- Luftman J, Zadeh HS (2011) Key information technology and management issues 2010–11: an international study. *Journal of Information Technology* 26(1):193–204
- Marshall P, Keahey K, Freeman T (2010) Elastic site: using clouds to elastically extend site resources. In: 10th IEEE/ACM international conference on cluster, cloud and grid computing (CCGRID), Melbourne
- Mateescu G, Gentsch W, Ribbens CJ (2011) Hybrid computing – where hpc meets grid and cloud computing. *Future Generation Computer Systems* 27(5):440–453
- MathWave Technologies (2010). EasyFit statistical package
- Mell P, Grance T (2009) The NIST definition of cloud computing. IITL National Institute of Standards and Technology
- Mohan TS (2011) Migrating into a cloud. In: Buyya R, Broberg J, Goscinski A (eds) *Cloud computing – principles and paradigms*. Wiley, Hoboken, pp 43–56
- Moore S (2011) Gartner survey of CIOs in Asia shows cloud computing tops the technology priority list as businesses focus on growth in 2011. Gartner Inc., Hong Kong. Press release
- Motahari-Nezhad HR, Stephenson B, Singhal S (2009) Outsourcing business to cloud computing services: opportunities and challenges. Working paper
- Oetiker T (2011) About RRdtool. <http://oss.oetiker.ch/rrdtool>. Accessed 2012-08-23
- Opitz A, König H, Szamlewska S (2008) What does grid computing cost? *Journal of Grid Computing* 6(4):385–397
- Petty C (2011) Gartner maps out the rapidly evolving market for cloud infrastructure as a service. Gartner Inc., Stamford, Press release
- Petty C, Tudor B (2010) Gartner says worldwide cloud services market to surpass \$68 billion in 2010. Gartner Inc., Stamford, Press release
- Raj P (2011) Enriching the ‘integration as a service paradigm’ for the cloud era. In: Buyya R, Broberg J, Goscinski A (eds) *Cloud computing – principles and paradigms*. Wiley, Hoboken, pp 57–96
- Rao AR, Hamed KH (1999) Flood frequency analysis. CRC Press, Boca Raton
- Rehman Z, Hussain FK, Hussain OK (2011) Towards multi-criteria cloud service selection. In: 5th international conference on innovative mobile and Internet services in ubiquitous computing, Seoul
- Reynolds E, Bess C (2009) Clearing up the cloud: adoption strategies for cloud computing. *Cutter IT Journal* 22(6/7):14–20
- Schwartz SC, Yeh YS (1982) On the distribution function and moments of power sums with log-normal components. *Bell Systems Technical Journal* 61(7):1441–1462
- Shroff G (2010) Enterprise cloud computing. Cambridge University Press, Cambridge
- Smith DM (2011) Hype cycle for cloud computing 2011. Gartner Inc., Stamford
- Staten J (2009a) Case study: USA.gov achieves cloud bursting efficiency using Terremark’s enterprise cloud. Forrester Research
- Staten J (2009b) Which cloud computing platform is right for you? Forrester research
- Strebel J, Stage A (2010) An economic decision model for business software application deployment on hybrid cloud environments. In: Schumann M, Kolbe LM, Breitenner MH, Frerichs A (eds) *Multikonferenz Wirtschaftsinformatik 2010*, Göttingen, pp 47–48
- Vaquero LM, Rodero-Merino L, Caceres J, Lindner M (2009) A break in the clouds: towards a cloud definition. *ACM SIGCOMM Computer Communication Review* 39(1):50–55
- Vecchiola C, Chu X, Mattess M, Buyya R (2011) Aneka – integration of private and public clouds. In: Buyya R, Broberg J, Goscinski A (eds) *Cloud computing – principles and paradigms*. Wiley, Hoboken, pp 251–274
- Walker E (2009) The real cost of a CPU hour. *IEEE Computer* 42(4):35–41
- Walker E, Briskin W, Romney J (2010) To lease or not to lease from storage clouds. *IEEE Computer* 43(4):44–50
- Weekes PA, Chambers JC, Mullick SK (1969) Lease-buy planning decisions. *Management Science* 15(6):B295–B307
- Weinhardt C, Anandasivam A, Blau B, Borissov N, Meinel T, Michalk W, Stößer J (2009) Cloud-computing – a classification, business models, and research directions. *Business & Information Systems Engineering* 1(5):391–399
- Yang H, Tate M (2009) Where are we at with cloud computing? A descriptive literature review. In: 20th Australasian conference on information systems, Melbourne
- Yang Z (2011) Disrupt the disruptor: a theoretical approach of cloud computing on IT outsourcing industry disruption. In: IEEE 9th international conference on dependable, autonomic and secure computing (DASC), Sydney