Information Systems 38 (2013) 495–508

Contents lists available at SciVerse ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

# Cost effectiveness of commercial computing clouds

Slaven Brumec [a], Neven Vrček [b,*]

[a] Koris d.o.o., Zagreb 10000, Croatia
[b] University of Zagreb, Faculty of Organization and Informatics, Varaždin 42000, Croatia

ARTICLE INFO

*Article history:*
Received 4 September 2012
Received in revised form
6 November 2012
Accepted 8 November 2012
Available online 28 November 2012

*Keywords:*
Cloud computing
Cost effectiveness
Fractional factorial design
Symmetric mediation plan

ABSTRACT

This paper presents the procedure for comparing costs of leasing IT resources in a commercial computing cloud against those incurred in using on-premise resources. The procedure starts with calculating the number of computers as depending on parameters that describe application's features and execution conditions. By measuring required execution time for different parameter values, we determined that this dependence is a second-order polynomial. Polynomial coefficients were calculated by processing the results of fractional factorial design. On that basis we calculated costs of computing and storage resources required for the application to run. The same calculation model can be applied to both a personal user and a cloud provider. The results will differ because of different hardware exploitation levels and the economy of scale effects. Such calculation enables cloud providers to determine marginal costs in their services' price, and allows users to calculate costs they would incur by executing the same application using their own resources.

Leasing in cloud establishes a business relationship: buyer wants to reduce costs, and cloud provider wants to generate profit. This relationship will be realized if the buyer and the provider agree on a mutually acceptable fair price that can be determined by the symmetric mediation plan.

All the steps in this procedure are integrated into CCCE method and represented as a process model.



## 1. Introduction

Although conceived as a technology a long time ago, it is over the last few years that cloud computing has established a more extensive commercial presence [3]. Most researchers [12,18,20], users and professional public [4,7,8,10] define commercial computer clouds as a set of network services intended for providing IT services, wherein their usage:

- can be characterised as being performed on the on-demand self-service basis,
- is available through standard network technologies and protocols,
- is based on resource virtualization,
- enables rapid elasticity, that is, scalability of provided services and available resources in accordance with the user's current business requirements, and
- is charged on the pay-as-you-go basis, like conventional utilities.

A number of authors have emphasised the reduction of capital costs as a key economic benefit of cloud computing in comparison with using conventional IT resources within private server farms [9,1,15,13,14]. The pay-as-you-go model, elasticity and scalability as well as the high level of virtualization of cloud IT resources enable the cloud user to avoid investing heavily into the acquisition of their own computing equipment in advance. Capital investment costs can be distributed into gradual and monitored operating costs of lease in the computing cloud. Furthermore, virtual

* Corresponding author. Tel.: +385 99 3390 827.
*E-mail address:* neven.vrcek@foi.hr (N. Vrček).

0306-4379/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.is.2012.11.002

computers are leased in the cloud include systems and applications software are incorporated into a configured server farm. The cloud user puts the cloud provider in charge of maintaining the leased virtual computers and software as well as of configuring the server farm, thus freeing their IT staff from these jobs and enabling them to focus on the development of IT innovations and business applications support.

On their portals, commercial cloud providers offer a calculator for estimating the costs of leased cloud IT resources which, however, fail to address two crucial questions: what is the amount of IT resources that a particular application will use? and What is the cost of maintenance of those applications of on-premise computers within a conventional server farm?

A good indicator of cost effective usage of IT resources in a commercial cloud, both for the buyer and the provider, is the fair price of leasing those resources. Let $C_P$[$/h] be the cost of computing resources, and $D_P$[$/GB h] the cost of storage resources that the buyer would incur in using their privately-owned resources to run the application. Let $C_S$ and $D_S$ be adequate costs for the cloud provider. The price $C$ and/or $D$ at which the cloud provider leases IT resources needs to be sufficiently high to ensure profit (i.e., $C > C_S$ and $D > D_S$), but also sufficiently low to make leasing cloud resources more cost effective for the buyer than purchasing and using on-premise computing resources (i.e. $C < C_P$ and $D < D_P$ need to be fulfilled). Although cloud providers argue that for the buyer leasing is a more favourable option than purchasing, in this paper we show that this does not necessarily apply to all cases. A prerequisite for objective decision-making regarding the purchase and lease of IT resources was setting the quantitative criteria, which was achieved by developing adequate cost calculation methods and their integration into a uniform method presented in the following sections of this paper. The aforementioned issues are scientifically relevant as they indicate possible future avenues of development of cloud computing business models as well as the impact of the cloud computing paradigm on business subjects and their competitive edge.

It has to be noted that in this paper commercial clouds are explored, which excludes academic-only computing clouds. Moreover, in this paper the security and legal issues of cloud computing usage are not specifically addressed. Instead, it is assumed that users act within a legal framework that provides an adequate environment for cloud computing and that cloud-related security concerns have been resolved by cloud providers.

## 2. Research overview

The data for the analysis of cloud services price as viewed from the providers' perspective is hard to find in the literature since investors do not tend to reveal their calculations to the public. That makes the design of appropriate cloud service models more difficult for scientists. An influential paper in the field of cloud cost analysis was contributed by Xinhui et al. [22], whose authors originate from the IBM Research Lab in China. Their results correspond with those in the study by Sawyer [16], which deals with the cost of data centres in general, not only cloud-based ones.

From the investors' point of view, an important feature of cloud computing is the total cost of ownership (TCO) over a cloud data centre. Once the TCO has been calculated, it is possible to define general factors that determine operating costs and the terms of cloud cost effectiveness, e.g. the minimum price of a CPU hour or other calculation factors. The paper by Xinhui et al. [22] is significant in that respect as it is applicable to the TCO calculation method for any computing cloud.

Important contribution to the analysis of cloud services price from the users' perspective was given by Walker [19]. Detailed description of the procedure for comparing the price of cloud computing resources against the cost of acquisition and maintenance of on-premises hardware is provided in [19], which is of critical importance in lease versus purchase decision-making.

Unlike conventional financial models for analysis of purchase or lease of capital assets, Walker's model does not only include amortisation, but also takes into consideration the change in the CPU performance over time in accordance with Moore's law. Furthermore, Walker's model makes it possible to calculate the price of a CPU hour at the moment of purchasing computing resources, the price of a CPU hour for the purchased resources including the regular annual computing capacity update and the theoretical price of a CPU hour for leasing computing resources in a cloud.

Leasing cloud computing resources can generally be observed as leasing processor time expressed by means of the price of a CPU hour. With certain cloud providers (e.g. Google App Engine) that is the only resource that is explicitly leased. With other providers (e.g. Amazon or Windows Azure) CPU time is leased implicitly, through virtual computers use, wherein configurations with precisely defined features such as the number of cores in a processor, disk capacity, size of RAM and bandwidth are available to users.

Walker's work is therefore universally applicable when the comparison of the price of on-premise resources against the price of leasing resources in any commercial cloud is concerned. Walker's research also proves the intuitive fact that the exploitation of cloud IT resources is the crucial factor in the price of computing services. A higher level of average processor occupation implies that the cloud provider can offer cloud services to the user at a lower commercial price.

## 3. Method for calculating cloud computing cost effectiveness (CCCE)

The review of the recent literature has led the authors of this paper to conclude that a comprehensive theoretical model that would link the application features to leased cloud resources has so far not been devised. Within our research we therefore developed a method for calculating cloud computing cost effectiveness (CCCE method), which can be implemented in deciding objectively whether to purchase or lease computing resources. The method is