Decision Support Systems 57 (2014) 320–330

Contents lists available at ScienceDirect

Decision Support Systems

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

Role of acquisition intervals in private and public cloud storage costs

Gabriella Laatikainen *, Oleksiy Mazhelis, Pasi Tyrväinen

Department of Computer Science and Information Systems University of Jyväskylä, Jyväskylä, Finland

ARTICLE INFO

Article history:
Received 26 November 2012
Received in revised form 17 September 2013
Accepted 22 September 2013
Available online 9 November 2013

Keywords:
Storage services
Cloud storage
Cost model
Acquisition interval

ABSTRACT

The volume of worldwide digital content has increased nine-fold within the last five years, and this immense growth is predicted to continue in the foreseeable future to reach 8 ZB by 2015. Traditionally, organizations proactively have built and managed their private storage facilities to cope with the growing demand for storage capacity. Recently, many organizations have instead welcomed the alternative of outsourcing their storage needs to the providers of public cloud storage services due to the proliferation of public cloud infrastructure offerings. The comparative cost-efficiency of these two alternatives depends on a number of factors, such as the prices of the public and private storage, the charging and the storage acquisition intervals, and the predictability of the demand for storage. In this paper, we study the relationship between the cost-efficiency of the private vs. public storage and the acquisition interval at which the organization re-assesses its storage needs and acquires additional private storage. The analysis in the paper suggests that for commonly encountered exponential growth of storage demand, shorter acquisition intervals increase the likelihood of less expensive private storage solutions compared with public cloud infrastructure. This phenomenon is also numerically illustrated in the paper using the storage needs encountered by a university back-up and archiving service as an example. Because the acquisition interval is determined by the organization's ability to foresee the growth of storage demand, via provisioning schedules of storage equipment providers, and internal practices of the organization, among other factors, organizations that own a private storage solution may want to control some of these factors to attain a shorter acquisition interval and thus make the private storage (more) cost-efficient.

1. Introduction

According to the IDC, the global volume of digital content has exhibited exponential growth and will grow from 1.8 ZB in 2011 to 2.7 ZB in 2012 and ultimately reach 35 ZB by 2020 [5,6]. As the volume of digital content grows, the global need for storage capacity rapidly increases, too.

To cope with the growing demand for storage, organizations may proactively build their private storage capacity or may alternatively opt for outsourcing their storage to the providers of public cloud infrastructure services, such as Amazon Simple Storage Service (S3), Box.com, and Apple iCloud. Decision makers consider several factors when they decide on possible adoption, such as cost, elasticity, data availability, security, data confidentiality and privacy, regulatory requirements, reliability, performance, integration with other services, personal preference, and added values [10]. However, cost considerations are perceived both as a risk and an opportunity, and the expected cost advantage is the strongest decisional factor that affects the perceived opportunities of IT executives [2].

If the organization decides to store its data in-house, it periodically estimates its future demand for storage and then proactively acquires and manages the storage infrastructure internally. Conversely, the use of cloud-based storage services gives the organization the flexibility to rapidly increase its storage capacity as the demand for storage grows, as well as the possibility to pay only for the volume of storage the organization actually uses within each charging period.

Because the cloud infrastructure capacity is usually paid for only when used, the cloud infrastructure providers include a so-called utility (pay-per-use) premium into their pricing [29]. As a result, the unit price per unit of time of a public cloud infrastructure capacity is usually more expensive compared with the unit cost of private capacity [29,9]. Still, if the demand for infrastructure services exhibits periodical or random peaks, the adoption of public cloud infrastructure is likely to offer cost advantages to organizations over the private infrastructure: this advantage is because the high premium charged by the public cloud provider is compensated by avoiding extensive periods of time when the private infrastructure would remain idle [29,12].

However, as opposed to the fluctuating demand for computing resources, the demand for storage often accumulates over time because newly created digital content only partially supersedes the already stored files. As a result, the use of public storage services may prove more expensive compared with the private solutions in the long term [27].

The cost-efficiency of public vs. private storage depends on a number of factors, such as the premium charged by the provider of public cloud infrastructure, the charging period (for the public storage) and

* Corresponding author. Tel.: +358 451444700.
E-mail addresses: gabriella.g.laatikainen@jyu.fi (G. Laatikainen), mazhelis@jyu.fi (O. Mazhelis), pasi.tyrvainen@jyu.fi (P. Tyrväinen).

0167-9236/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2013.09.020

the storage acquisition interval (for the private storage), the intensity of incurred data communications, the predictability of the growth of storage needs and the storage growth profile [27,30]. Due to the continuously increasing storage demand, the length of the acquisition intervals and growth predictability are among the most critical factors in storage cost analysis; however, they have to date not been studied in detail. Therefore, we study the effect of the private storage acquisition interval on the cost-efficiency of private vs. public storage in this paper. This interval can be determined by the organization's ability to foresee the growth in storage demand, via the provisioning schedule of the storage equipment provider, the internal practices of the organization, etc. The paper analytically shows that for commonly encountered exponential growth of storage demand, shorter intervals for which the organization re-assesses its storage needs and acquires additional storage increase the likelihood that a private storage solution is less expensive compared with the public cloud infrastructure. Numerical experiments are employed to illustrate this dependency using the storage needs encountered by a university back-up and archiving service as an example.

The analysis of storage costs in the paper focuses on the storage needs and their growth and predictability, storage acquisition interval, as well as the costs incurred due to the transfer of data to and from the storage location. The storage costs may also be affected by additional factors, such as the economies and diseconomies of scale, the cost of capital, the required level of availability and durability and the possibility to use provenance data. Combined, these and other factors are likely to have a complex, non-linear effect on the overall costs, which makes them difficult to analyze [12]. To simplify the analysis, these additional factors were assumed to either have a minor effect or similar effect on the costs of both the private and public storage solutions. Hence, these factors are outside of the scope of the paper.

The remainder of the paper is organized as follows. In the next section, the related works on the cost-efficient use of cloud infrastructure are reviewed. In Section 3, an analytical model for comparing the cost efficiency of private vs. public storage is introduced, in which the effect of the acquisition interval is taken into account. Numerical experiments that illustrate the effect of the acquisition interval and its interplay with various other factors are provided in Section 4. In Section 5, the practical implications and limitations of the obtained results are discussed. Finally, Section 6 summarizes the obtained results and outlines the directions for further work.

2. Related works

In recent years, extensive research efforts have been devoted to the cost-efficient use of cloud infrastructure services in general, and cloud storage services in particular. A short overview of the recent research in this domain is provided below.

A number of works have focused on the cost-efficient use of cloud infrastructure and the factors that affect it. In particular, the cost benefit of using cloud bursting, i.e., offloading the computing load during peak times to a public cloud infrastructure, has been analytically investigated in [29,28]. The cost efficient allocation of computing load to the private and the public portions of a hybrid cloud infrastructure was also studied in [8,3] and in [13,12], where the communication overheads were also considered. The cost-optimal time of using the public cloud has been shown to be the inverse of the premium charged by the public cloud provider assuming negligible data communication overheads.

The economies of scale, i.e., the decline in the cost per unit of a service with the number of units produced [22], may affect the cost-efficiency of private vs. public cloud infrastructure as well. These economies of scale are manifested e.g., in the volume discount offered for the cloud infrastructure capacity, and the cost of a hybrid cloud may exceed the cost of a private or a public cloud infrastructure in the presence of such discounts [12].

The cost-optimal allocation of individual computing tasks to private and public cloud resources was also approached as a multi-integer linear programming problem in [24]. Based on the results of a simulation study, the authors found little or no cost benefits in offloading the peaks of the workload, although the preliminary character of the study and the complex nature of the optimization model make it difficult to interpret the results. Walker [27] compared the acquisition and leasing of storage as alternative investment decisions based on their Net Present Value (NPV). The estimation of the NPV considers the dynamics of the demand for storage, the gradual decline of acquired and leased storage prices, the disk replacements due to possible disc failures, and the salvage value of the acquired discs at the end of their use time. Using numerical examples, the authors illustrated that leasing represents a cost-optimal alternative for small- and medium-sized enterprises, whereas acquiring storage is likely to be less expensive in the long term for large enterprises.

Mastroeni and Naldi [11] further revised Walker's model by replacing the deterministic estimation of the pricing dynamics and disc failure dynamics in [27] with probabilistic models. Based on these models, the authors arrive at a probabilistic distribution of differential NPV values and use its median to determine the economically justifiable alternative. Note that in both [27] and [11], the costs are accounted on a yearly basis; thus, the role of acquisition intervals shorter than a year is not visible in these models.

Uttamchandani et al. [26] introduced BRAHMA, a tool that applies constraint-based optimization to cost-optimally supply the storage demand with a mixture of in-house and public cloud storage resources. The tool suggests an optimal placement both for the storage and for the system administrators based on customer storage needs and the projected growth thereof over a look-ahead period, as well as associated service level objectives. The tool helps to identify the optimal sourcing if the customer and the storage service provider have a heterogeneous set of devices and human resources that have different costs. However, to the best of our knowledge, the tool assumes a perfect knowledge of the customer demand growth and fails to consider the storage acquisition intervals; as a result, the cost of the over-provisioned storage is not visible when using the tool.

Constraint-based optimization has also been employed by Trummer et al. [25] to optimally allocate applications to the cloud along with their storage resources. The authors' approach assumes that the resource requirements are known in advance, which is similar to the BRAHMA tool. The effects of imperfect knowledge and resulting storage over-provisioning are not considered.

In addition to acquiring storage capacities, organizations may maximize the cost-efficiency of cloud solutions by storing only the provenance for data and regenerating the rest when needed [1]. Yuan et al. proposed different strategies to find the best trade-off of storage and computational costs by storing the appropriate intermediate data in cloud storage [31,33,32]. Muniswamy-Reddy et al. emphasized the need for incorporating provenance services in cloud storage providers, analyzed several alternative implementations to collect provenance data, and use the cloud as a backend [17,16,15].

Finally, Weinman [30] considered the delay with which the required resource is provisioned and analyzed both the cost of over-provisioning (i.e., unused resources) and under-provisioning (i.e., the opportunity cost of unserved demand). The author discussed the role of provisioning time given a possibility to predict the future demand over a specific forecast visibility; however, the paper only considered cases with a zero forecasting visibility.

In summary, while a number of works have focused on the cost-efficient use of private and public infrastructure resources, relatively little attention has been devoted to the role of the acquisition intervals in the cost-efficient use of private vs. public storage capacity. Therefore, a storage cost model is introduced below in which the effect of the acquisition interval is taken into account.