



# InterPARES Trust Project

## Research Report

<b>Title:</b>	Economic models for storage of records in the cloud (StaaS) – A critical review of the literature (EU18)
<b>Status:</b>	Draft (restricted)
<b>Version:</b>	1.0
<b>Date submitted:</b>	
<b>Last reviewed:</b>	
<b>Author:</b>	InterPARES Trust Project
<b>Writer(s):</b>	Prof Julie McLeod, iSchool, Northumbria University Ms Brianna Gormly, GRA, UBC
<b>Research domain:</b>	Resources cross domain
<b>URL:</b>	

## Document Control

Version history			
Version	Date	By	Version notes
V0	27/04/2015	JM, BG	Initial draft
V1	10/07/2015	JM, BG	Completed report

## **Introduction**

The aim of this project was to provide a critical review and analysis of published literature that either examines economic models for storing records in the cloud (storage as a service - StaaS) or provides in-depth discussion and/or analysis of the pricing and costing of StaaS. The rationale for the review was twofold:

- A body of literature exists, much of it from service providers and consultants/consultancy companies, which highlights the financial/economic benefits of using cloud services for the storage of digital information. However, on what basis are these claims made and, in particular, what are the underpinning economic models?
- There is evidence that archives/records professionals (ARM) are increasingly using the cloud for storing collections of digital records. As information professionals assess moving the storage of some/all of their records/archival collections to the cloud, what economic models are available to them for estimating the cost as well as the medium to long-term financial implications for their organisations?

The objectives were:

1. To identify any economic models for using cloud for StaaS
2. To compare the models in terms of their underpinning theory and assumptions
3. To identify and evaluate any case examples where these models, or other approaches, have been explicitly used to support decisions on using cloud or StaaS.

The review focused on identifying economic models for estimating the cost of medium to long-term storage of records/digital information in the cloud; the focus was *not* on the cost of using the cloud for digital preservation per se.

## **Literature review method**

The aim was to comprehensively search relevant secondary sources (abstracting and indexing databases) to identify primary literature on the topic. A purposive selection of secondary sources and the websites of relevant organisations, encompassing a cross-disciplinary perspective, were searched. The key disciplines were: archives and records management; computer science and IT; and business. Details, including the search strategy and results, are included in the appendix which also documents how sources were identified for inclusion in the annotated bibliography. One additional source was the product of a previous InterPARES study (InterPARES 3, General Study 16, 2013) on the costs and benefits of digital preservation.

Sources in the annotated bibliography are divided into three tiers, reflecting their relevance to the project:

- Tier 1 comprises the sources most directly relevant to the project's research question. The majority present models for determining the cost of cloud storage, and the others also provide discussions pertinent to the economics of using the cloud for storage;
- Tier 2 includes sources which model or discuss the overall costs of cloud computing. Many of these deal with storage as one of many cost considerations. Tier 2 also includes relevant studies on the costs of archival storage;
- Tier 3 consists of research relevant to understanding the costs of cloud computing. These sources deal with very specific situations (e.g. science data, hybrid clouds) or present other aspects of cost modeling of limited relevance to this project. They are included for potentially broader interest and value.

### **Summary critical narrative**

This narrative is based on Tier 1 sources only, being the most relevant to the project objectives.

### ***Economic models for using cloud for StaaS***

The most relevant work on economic models for cloud StaaS is that of Walker, Brisken and Romney (2010); Naldi and Mastroeni (2011; 2013; 2014); Mazhelis, Fazekas and Tyrväinen (2012) and Laatikainen, Mazhelis and Tyrväinen (2014); Rosenthal and Vargas (2012); DC Rosenthal et al. (2012) and DSH Rosenthal et al. (2012). Parts of the 4C (Collaboration to Clarify the Costs of Curation) project (2013-15) are also relevant, in particular the 'Evaluation of Cost Models and Needs & Gap Analysis' and the related summaries of the 10 cost models. Other scholars have contributed relevant work to the topic, viz. Khajeh-Hosseini et al. (2011); Wang et al. (2012); and Dutta and Hasan (2013). Reichman (2011) presents the work of consultancy company Forrester Research Inc.

Walker, Brisken and Romney's (2010) work is the earliest work found, which Naldi and Mastroeni (2011; 2013; 2014) cite and explicitly build upon to address its perceived weaknesses. This body of work is cited by Mazhelis, Fazekas and Tyrväinen (2012) and by Laatikainen, Mazhelis and Tyrväinen (2014), although their model is different. Khajeh-Hosseini et al. (2011) and Wang et al. (2012) all cite Walker and colleagues. These authors are all situated in the computer science / information systems discipline (with Mastroeni in economics) and publish in that field. In contrast, the work of DC and DSH Rosenthal and colleagues and the 4C project team, whose summary of cost

models includes the Rosenthals' economic model for long-term storage, is situated in the library/archives discipline and focuses on digital preservation.

With the exception of Dutta and Hasan, who are based in computing and information sciences and cite DSH Rosenthal et al., there is little citation between the two disciplines (i.e. computer science and archival science). It appears that complementary work has been undertaken in parallel 'silos'. If this separation in the scholarship plays out in practice then there is a danger that records/archives professionals may not be cognisant of and consider the literature from computer science and/or discuss the economics of cloud storage with their computer science/IT colleagues. Conversely, scholars in computer science may not be cognisant of the needs or concerns of records/archives professionals in this space.

### ***Underpinning theory and assumptions***

The following economic or financial/management accounting theories underpin the models presented in the work of these authors:

1. Discounted Cash Flow including Net Present Value, Differential Net Present Value and Internal Rate of Return  
*(Naldi and Mastroeni; Walker, Brisken and Romney; Wang et al.; Khajeh-Hosseini et al. See also: Rosenthal and Vargas; DC and DSH Rosenthal et al.)*
2. Monte Carlo models and Kryder's Law  
*(DC and DSH Rosenthal et al.; Rosenthal and Vargas)*
3. Full Cost Accounting including Total Cost of Ownership  
*(Dutta and Hasan; Reichman)*
4. Acquisition interval - length of acquisition of additional storage  
*(Laatikainen, Mazhelis and Tyrväinen; Mazhelis, Fazekas and Tyrväinen)*

#### ***1. Discounted Cash Flow including Net Present Value, Differential Net Present Value and Internal Rate of Return***

Discounted Cash Flow (DCF) techniques are based on the principle of the value of money (spent or invested) over time; i.e. a unit of money today having a different (usually lower) value in the future, taking account of inflation, interest rate (the discount rate) and returns. Although 'standard' economic techniques, they are sometimes criticised because they assume the interest rate is constant rather than variable over time, as in practice. In the context of modelling digital storage costs over longer periods, they are potentially less useful.

Net Present Value (NPV) is the sum of the present values of all the cash flows relating to a project or investment, i.e. cash inflows (cash earned) and cash outflows (cash spent). A positive NPV indicates a profit, a negative NPV a loss. In a buy or lease scenario if the NPV(buy) is greater than the NPV(lease) then the decision should be to buy. The Differential Net Present Value (DNPV) considers the difference between the two NPVs, rather than their absolutes, and is easier to calculate<sup>1</sup>. Both models enable a comparison of the cost of purchasing vs leasing assets for some purpose – in this case for digital storage. They each consider a number of factors; for example, capital costs (e.g. purchase, interest rate), operating costs (e.g. energy, personnel), disc price trends, disc replacement rates and hardware salvage value.

In their use of the DNPV model, Walker, Brisken and Romney (2010) use past data to predict the future cost of factors, such as disc price and salvage value, as well as future disc replacement rates. Naldi and Mastroeni (2011, p1) perceive this to be “deterministic” and a weakness of the model. Their DNPV model is a more sophisticated probabilistic model that takes into account future “unknown” or “random” changes in, for example, leasing price and disc failure, and also incorporates risk measures (Naldi and Mastroeni, 2011, p1 and 2013, 2014). This results in the DNPV becoming a range of values.

Wang et al. (2012) use the NPV model but address its perceived weakness by incorporating the Internal Rate of Return (IRR), i.e. the interest rate required for the NPV to be zero. They also include the concept of a ‘burstiness filter’ (referring to “cloud bursting” when a peak computing level is hit and data is transferred from the data centre to the cloud) to detect what work might be transferred to the cloud to increase cost savings.

Khajeh-Hosseini et al. (2011) also use NPV to compare the financial cost of storage options as part of their Cloud Adoption Toolkit. However, the toolkit is not a predominantly financial one; rather it comprises five decision-making tools, many of which are soft (qualitative) and were under development at the time of publication. The ‘technology suitability analysis’ is a short checklist of questions; the ‘energy consumption analysis’ uses performance per unit of energy characteristics of physical machines and performance requirements of virtual machines; the ‘stakeholder impact analysis’ considers the socio-political benefits and risks (of the cloud); the ‘responsibility modeling’ identifies who is responsible for what and determines “the practical, social, and political viability “ of their discharge to meet requirements; and the ‘cost modeling’

---

<sup>1</sup> An element of estimated profit is included in calculating NPV for both buy and lease scenarios. These cancel each other out in a DNPV, removing the need to estimate and making the calculation easier.

tool uses UML deployment diagrams to provide estimates of the operational costs of a system. The financial cost model used is NPV.

## 2. Full cost accounting

Full cost accounting recognises a wider range of costs than standard cash flow (financial) methods, for example economic, social and environmental costs. Total Cost of Ownership (TCO) can be used in full cost accounting and, as the name suggests, is the sum of all expenditures of a project or system (e.g. power, personnel/labour, hardware). These account for direct and indirect costs, including overheads, but do not account for the time value of money.

Dutta and Hasan's (2013) full cost accounting model includes initial infrastructure set up costs, floor rent, energy, service (e.g. software development, hardware repair), disc disposal and environmental costs (e.g. carbon emissions). The two Rosenthal et al.'s (2012; 2012) first model is based on TCO and includes the cost of energy, labour, infrastructure and disc replacement. The latter takes account of Kryder's Law<sup>2</sup>, which states that storage density of discs doubles every two years, though it is widely translated into the exponential decrease in digital storage cost. (See later discussion of Monte Carlo methods).

Reichman (2011) suggests TCO is a good approach but difficult to use accurately in practice. Instead "a more pragmatic approach is to compare only the costs that change between the two scenarios, known as relative cost of operations." (p13) On this basis, the (changing) the factors included in the (Forrester) model he presents are: service life of storage; storage acquisition cost; redundancy copies; storage utilisation; personnel; infrastructure cost (facilities and energy); maintenance and data migration. The model is built into a spreadsheet tool making it possible to calculate the annual storage cost of internal vs cloud storage.

## 3. Monte Carlo models and Kryder's Law

The two Rosenthal et al.'s (2012; 2012) second model attempts to address the "inadequacy" of NPV models based on DCF, since the assumption of a constant interest rate is potentially more problematic for longer term modeling of storage costs. A Monte Carlo (probability) model calculates the endowment (i.e. amount of money) that must be invested with the data to cover the cost of storage over its anticipated lifetime. This

---

<sup>2</sup> Walter, C. (2005). Kryder's Law. *Scientific American*, August, 293, p.32-33 and Wikipedia entry [https://en.wikipedia.org/wiki/Mark\\_Kryder#Kryder.27s\\_Law](https://en.wikipedia.org/wiki/Mark_Kryder#Kryder.27s_Law)

involves “projecting both the Kryder’s Law decrease in cost and future interest rate which will apply to the unexpended parts of the endowment” (DSH Rosenthal et al., 2012, p3). Rosenthal and Vargas (2012) also refer to Monte Carlo models and consider Kryder’s Law in assessing whether or not storing LOCKSS<sup>3</sup> boxes in the cloud would be better financially and discover that cloud storage pricing has not decreased according to this law.

#### 4. *Acquisition intervals for additional storage*

Laatikainen, Mazhelis and Tyrväinen (2014) and Mazhelis, Fazekas, and Tyrväinen (2012) propose a storage cost model which incorporates the length between intervals at which an organisation evaluates its storage needs (including predicting demand for storage - “growth predictability”) and acquires additional in-house storage. They consider these to be “critical factors in storage cost analysis” that are not accounted for in the DCF/NPV models (above). They suggest “the cost efficiency of private vs. public storage depends on the price difference of the private and public storage, the interval at which the storage can be acquired, and the accuracy with which the future needs for the storage can be predicted” (Laatick, Mazhelis and Tyrväinen, 2014, p 322). Their mathematical model therefore incorporates storage capacity/demand, the unit cost of storage, depreciation and redundancy and considers the effect of the length of the acquisition interval on the cost of private storage.

### **Case examples of the practical application of the models to support decisions on using cloud StaaS**

The following table, organised by the underpinning economic or financial model, summarises the case examples and scenarios presented in the sources discussed above. There is a limited number of ‘real’ case examples and none in a records/archives specific context. The ‘real’ case examples are all based on university departments or services; the hypothetical scenarios all consider the relative size of the organisation (small, medium, large) irrespective of sector, and make comparisons. Most examples include a caveat(s), for example additional factors to consider or risk assessment, in drawing conclusions.

---

<sup>3</sup> Lots of Copies Keep Stuff Safe™, Stanford University



Case example/Scenario	Conclusion (buy vs lease; cloud vs in-house)	Authors
<b>1. Discounted Cash Flow including Net Present Value, Differential Net Present Value and Internal Rate of Return</b> See also (3): <i>Rosenthal and Vargas; DC and DSH Rosenthal et al.</i> )		
Three scenarios - a single-user computer, medium-size and large organisations	Single-user computer: cloud is more cost-effective for storage of less than four years, purchase is recommended for long-term storage. Medium-size organisations: cloud storage is the best option. Large organisations (over 1000 servers): cloud is cost-effective for up to nine years, purchase is more economical for longer periods	Walker, Brisken, and Romney (2010)
Pricing comparison of providers: Dropbox, SugarSync, IDrive, Google Drive, Carbonite, Symform, Mozy, Amazon	Two types of pricing: tapering pricing (“declining block rate charge”) and bundling (or “quantity discount”). All providers surveyed, except Amazon, use bundling. For individual users, Google Drive and IDrive are the best options. For businesses, Amazon, Carbonite, and SugarSync appear to offer the best plans, though Dropbox is also a contender when factors beyond the calculation are considered.	Naldi & Mastroeni (2013)
Hypothetical ‘simulation scenario’ with values for each parameter in the model (based on current costs etc) and size of company (small, large)	Larger companies looking toward storage over the long term (10+ years) will benefit from a buy decision. For smaller companies the cost of an insurance policy to protect against the risk of making the wrong decision is “affordable” over the shorter time according to their pricing formula.	Naldi & Mastroeni (2014)
Scenarios: medium and large companies	Cloud storage has greater cost benefits for medium-sized than for large-sized companies and for long-term rather than short-term investment. They assess the risk of making the wrong lease-or-buy decision, employing the Value-at-Risk risk measure, to demonstrate that risk is greatest when the possible profits from both buy and lease decisions are close to equal.	Mastroeni & Naldi (2011)
Hypothetical scenarios using Amazon S3	Cloud storage is more beneficial for small organisations (annual storage growth rate of 1TB) than for large organisations (annual storage growth rate	Wang et al. (2012)

(advertised pricing)	of 10 TB)	
School of Computer Science University of St. Andrews, UK compared with Amazon Web Services (AWS)	Little difference between in-house (buying servers) and cloud storage costs, but need to consider factors beyond financial considerations, such as organisational change.	Khajeh-Hoss eini et al. (2012)
<b>2. Monte Carlo models and Kryder's Law</b>		
Hypothetical scenarios with values for each parameter in the short and long term models based on current costs	Pricing models of cloud storage services (at the time) mean cloud is not an appropriate option for long-term storage.	DSH Rosenthal et al. (2012)
LOCKSS boxes using Amazon S3 (simple storage service)	Local disk storage is cheaper for long-term storage because cloud storage pricing has not decreased according to Kryder's Law. Proposed adjustments to Amazon S3, to make the service more cost effective, are vindicated by the introduction of Amazon Glacier, but, if Glacier pricing follows that of S3, it will be a more expensive long-term option than local storage.	Rosenthal & Vargas (2012)
<b>3. Full Cost Accounting including Total Cost of Ownership</b>		
Computer and Information Science Department Data Centre (small), University of Alabama Birmingham, USA compared with Amazon S3 (advertised pricing)	<p>Cost for storing 1 byte per year:</p> <ul style="list-style-type: none"> <li>• in house = <math>71.51 \times 10^3</math> picocents</li> <li>• Amazon S3 = <math>88.37 \times 10^3</math> picocents</li> </ul> <p>(1 US picocent = <math>\\$1 \times 10^{-14}</math>)</p> <p>The costs are relatively similar but factors such as pricing, scale of operation and data redundancy should be considered.</p>	Dutta and Hasn (2013)
Generalised comparison between internal and cloud storage of 100 TB of data	Demonstrates a 74 percent reduction in cost with the cloud but cautions about other issues involved in using the cloud.	Reichman (2011)

4. Acquisition interval - length of acquisition of additional storage		
<p>Oxford University Computing Services compared with Amazon S3 (advertised pricing)</p>	<p>In a typical case of exponentially growing storage demand (public) cloud storage is more cost effective when the intervals between assessments of private storage are longer. However, the acquisition interval at which private storage becomes more financially beneficial is also affected by other factors: the utility premium charged by the cloud provider, necessary storage redundancy, and the costs of transferring data to and from the cloud.</p>	<p>Laatikainen, Mazhelis &amp; Tyrväinen (2014); Mazhelis, Fazekas, &amp; Tyrväinen (2012)</p>

## **Issues of Trust**

Putting data/records in the hands of a service provider raises issues of trust – trust that it is safe, accessible and can be transferred/extracted. Such issues have been the subject of research and are central to the work of InterPARES Trust. However, what of trust in making the appropriate decision and/or business case to use (or not) the cloud for the medium to long-term storage of digital information? What of trust in a service provider to offer, and continue to deliver, an economically viable and sustainable storage service? The economic models discussed above either implicitly or explicitly raise and/or attempt to address issues of trust in the use of cloud storage for digital information. For example, uncertainty in future costs (Mastroeni and Naldi, 2011), and the failure of cloud service providers to reduce prices at the same rate as Kryder's law, hence increasing their profit (DSH Rosenthal et al., 2012).

## **Next steps/Future work**

This literature review provides a set of resources and an analysis of potential economic models, their theoretical basis and underlying assumptions, their application in 'real' or hypothetical scenarios and issues of trust that they raise. ARM professionals might use such models to inform their decision-making and business case preparation for using (or not) cloud storage services.

Since no published case examples of the use of any economic models by ARM professionals were identified, a proposal will be submitted to conduct an empirical project exploring the use or otherwise of such models in practice to estimate/predict the long-term economic implications of moving to the cloud for SaaS. This project will potentially offer real or hypothetical ARM case examples which could lead to recommendations on the use of economic models for decision-making in the context of cloud storage for digital information. This issue is key to trusting in the economic viability and sustainability of using cloud storage for digital information (not specifically digital preservation).

## Appendix

This appendix demonstrates the manner in which sources were identified for inclusion in the annotated bibliography. One additional source was the product of a previous InterPARES study (InterPARES 3, General Study 16).

### Databases

Database	Successful search terms	Relevant sources <sup>4</sup>	Notes
LISTA	cloud computing AND cost model	2	One source was found indirectly through a review.
LISA	cloud computing AND cost model	1	
IEEE Xplore	cloud computing AND economic model; cloud computing AND pricing model; cloud computing AND costing model	5	Cloud computing AND pricing model and cloud computing AND costing model returned 700-900 results, so the results list was examined only to the point at which it ceased to present results directly relevant to the query.
ACM	cloud computing AND economic model; cloud computing AND cost model; cloud storage AND pricing; cloud storage cost	9	Searches were limited to abstracts only.
Google scholar	n/a	9	Google Scholar initially identified DSH Rosenthal et al. "The Economics of Long-Term Digital Storage" and was subsequently used to identify sources citing that and also Walker, Brisken, and Romney "To Lease or Not to Lease from Storage Clouds."
Business Source Complete	cloud computing AND cost model	4	One source was found indirectly through a review.
ABI/Inform Global	cloud computing AND cost model	3	Searches were limited to abstracts only.

<sup>4</sup> This number includes related papers referenced in the "notes" field of entries in the bibliography that were identified in the search. Other related work included in this field was sometimes identified through references within the articles themselves.

*Note*

The secondary source databases covered the literature as follows:

- LISA/LISTA - archives and records management
- IEEE Xplore and ACM - computer science/information technology
- Business Source Complete and ABI/Inform Global - business, financial and economic
- Google Scholar – identification of current/key research actors and activity, and citations to specific references found

### Websites

<b>Websites</b>	<b>Relevant sources</b>	<b>Notes</b>
Professional organisations (SNIA, SAA, ACA, ARMA, ARA, IRMS, RIM PA, ICA, CSA)	3	
National Archives (TNA, NARA, LAC, NAA)	3	Two sources were found indirectly through references in a report.
Major Organisations/Consultancies (AIIIM, Gartner, Forrester)	2	
Cloud Service Providers	2	The search focused on the websites of cloud service providers included in a list provided in the TNA report, <i>Guidance on Cloud Storage and Digital Preservation: How Cloud Storage can address the needs of public archives in the UK</i> .
Collaboration to Clarify the Costs of Curation (4C)	5	This category includes the annotation for the 4C website and other sources for which it provided external links.

## **Annotated bibliography**

Sources in the annotated bibliography have been divided into three tiers, reflecting their different levels of relevance to the project. Tier 1 sources pertain directly to economic models for SaaS. These include the most useful sources described in the narrative summary. To recognise the large body of research on the overall costs of cloud computing, a second tier has been included. Many Tier 2 sources deal with storage as one of many cost considerations in the decision to move to the cloud. Other sources were identified in the search process that pertain to modeling the costs of cloud computing but which concern very specific situations less relevant focus of this project (e.g. studies specifically dealing with hybrid clouds, science data, etc.). These sources were retained as Tier 3.

## Tier 1

Tier 1 comprises the sources most directly relevant to the project's research question. The majority of these sources present models for determining the cost of cloud storage, and the others also provide discussions pertinent to the economics of using the cloud for storage.

### **Title: Collaboration to Clarify the Costs of Curation (4C)**

**Date:** 2015

**Source type:** Website

Collaboration to Clarify the Costs of Curation (4C) is a European Commission funded project with the goal of promoting better investment in digital curation. Unlike other projects which have purely examined cost, 4C also considers the benefits to be gained through curation. Moreover, though significant work has been done in the area of cost analysis, relevant stakeholders do not necessarily know that this work exists or how to apply it. Several project deliverables focus on examining and explaining existing work. These deliverables include the project's "Evaluation of Cost Models and Needs & Gap Analysis" and related summaries of the ten cost models. 4C was a 24-month project which ended in January 2015. The project culminated in a roadmap explaining the steps to be taken in the next five years (2015-2020), presented to the European Commission.

**Full citation:** 4C: Collaboration to Clarify the Costs of Curation. (2015). Retrieved from <http://www.4cproject.eu/>.

**Notes:** The model most relevant to cloud storage is the Economic Model for Long-Term Storage (see below annotations for DC Rosenthal et al. and DSH Rosenthal et al.). Though dealing with the range of costs in the records' lifecycle, LIFE<sup>3</sup> Costing Model and Total Cost of Preservation also make points relevant to cloud storage (see below tier 2 annotations). The project Cost Model for Digital Preservation also provides a resource relevant to modeling the cost of archival storage more generally (see tier 2 annotation for Nielsen, Thirifays, and Kejser).

**Authors:** Dutta, A. K., & Hasan, R.

**Title:** "How Much Does Storage Really Cost? Towards a Full Cost Accounting Model for Data Storage"

**Date:** 2013

**Source type:** Conference paper

The authors apply a full cost accounting model to the economics of cloud storage. Their model addresses the tendency to view storage cost as purely the cost of storage material. The authors, instead, consider a range of factors contributing to cost, specifically: "initial cost," "floor rent," "energy," "service," "disposal cost," and "environmental cost" (pp. 32-33). Using the sum of these costs, they can calculate that total price of storing one byte for a year. The authors apply their model in a case study using the data center at the Computer and Information Science department at the University of Alabama Birmingham and compare their findings to the advertised pricing for Amazon S3.



**Full Citation:** Dutta, A. K., & Hasan, R. (2013, September 18-20). How much does storage really cost? Towards a full cost accounting model for data storage. In *Economics of Grids, Clouds, Systems, and Services*. Paper presented at the 10th International Conference: GECON 2013, Zaragoza, Spain (pp. 29-43). Springer International Publishing. doi: 10.1007/978-3-319-02414-1\_3.

**Notes:** Dutta and Hasan's literature review includes DSH Rosenthal et al. "The Economics of Long-Term Digital Storage" (see below annotation).

**Authors:** Khajeh-Hosseini, A., Greenwood, D., Smith, J. W., & Sommerville, I.

**Title:** "The Cloud Adoption Toolkit: Supporting Cloud Adoption Decision in the Enterprise"

**Date:** 2011

**Source type:** Journal article

The article presents the Cloud Adoption Toolkit, an aide for decision-makers considering cloud services. It involves a combination of tools for addressing a range of relevant factors: "Technology Suitability Analysis, Energy Consumption Analysis, Stakeholder Impact Analysis, Responsibility Modeling, and Cost Modeling" (p. 463). The authors apply their work in a case study at the University of St. Andrews School of Computer Science. This study showcases the "Cost Modeling" tool (the most developed component of the toolkit). The authors conclude that the School of Computer Science should buy servers if funding is available. If the necessary up-front funding is not available, the School of Computer Science could consider a cloud provider; however, the School would need to restructure its IT architecture to work within the elastic nature of the cloud or risk incurring substantial costs. Moreover, the authors stress the importance for models to consider factors beyond purely financial considerations, such as organisational change.

**Full citation:** Khajeh- Hosseini, A., Greenwood, D., Smith, J. W., & Sommerville, I. (2012). The cloud adoption toolkit: supporting cloud adoption decisions in the enterprise. *Software: Practice and Experience*, 42(4), 447-465. doi:10.1002/spe.1072.

**Notes:** The authors reference Walker, Brisken, and Romeny (see below annotation). For another article discussing the Cloud Adoption Toolkit, including the cost modeling tool, see:

Khajeh-Hosseini, A., Sommerville, I., Bogaerts, J., & Teregowda, P. (2011, July 4-9). Decision support tools for cloud migration in the enterprise. Paper presented at 2011 IEEE International Conference on Cloud Computing (CLOUD), Washington, D.C. doi: 10.1109/CLOUD.2011.59.

**Authors:** Laatikainen, G., Mazhelis, O., & Tyrväinen, P.

**Title:** "Role of Acquisition Intervals in Private and Public Cloud Storage Costs"

**Date:** 2014

**Source type:** Journal article

The article demonstrates that the length between intervals at which the organisation evaluates its storage needs and acquires additional storage affects the cost benefits of in-house vs. cloud storage. In a typical case of exponentially growing storage demand,

shorter intervals between evaluations increase the likelihood that a private storage option will be more cost effective than using a public cloud. The authors also consider linear and logarithmic growth and assess the affect of data transfer (to storage and back to the user) on the calculated cost. The model is applied in an example examining the storage demand at Oxford University Computing Services.

**Full citation:** Laatikainen, G., Mazhelis, O., & Tyrväinen, P. (2014). Role of acquisition intervals in private and public cloud storage costs. *Decision Support Systems*, 57, 320-330. doi:10.1016/j.dss.2013.09.020.

**Notes:** The authors' literature review references Walker, Brisken, and Romney (see below annotation) and Mastroeni and Naldi "Long-Range Evaluation of Risk in the Migration to the Cloud" (see notes for Mastroeni and Naldi "Storage Buy or Lease Decisions Under Price Uncertainty"). Elsewhere, they also cite Khajeh- Hosseini et al. (see above annotation).

**Authors: Mastroeni, L. & Naldi, M.**

**Title: "Storage Buy-or-Lease Decisions in Cloud Computing Under Price Uncertainty"**

**Date:** 2011

**Source type:** Conference paper

Mastroeni and Naldi build on of the work of Walker, Brisken, and Romney (see below annotation), presenting a model to assist companies making the decision to buy disc storage or lease storage in the cloud. Like the earlier model prosed by Walker, Brisken, and Romney, the authors calculate the Net Present Value of the projected storage in terms of in terms of capital expenses, operational expenses, and salvage value (i.e. reselling unused discs). However, they adapt their model to account for leasing price fluctuation and disc failure. The authors simulate the application of their model for medium and large companies, demonstrating that cloud storage has greater cost benefits for medium-sized rather than large-sized companies and for long-term rather than short-term investment. They go on to assess the risk of making the wrong lease-or-buy decision, employing the Value-at-Risk risk measure to demonstrate that risk is greatest when the possible profits from both buy and lease decisions are close to equal.

**Full citation:** Mastroeni, L. & Naldi, M. (2011, June 27-29). Storage buy-or-lease decisions in cloud computing under price uncertainty. Paper presented at 2011 7th EURO-NGI Conference on Next Generation Internet (NGI), Kaiserslautern. doi: 10.1109/NGI.2011.5985868.

**Notes:** For another paper on this project see:

Mastroeni, L. & Naldi, M. (2011, September 5-7). Long-range evaluation of risk in the migration to cloud storage. Paper presented at 2011 IEEE 13<sup>th</sup> Conference on Commerce and Enterprise Computing (CEC), Luxembourg. doi:10.1109/CEC.2011.47.

**Author: Mazhelis, O., Fazekas, G., & Tyrväinen, P.**

**Title: “Impact of Storage Acquisition Intervals on the Cost-Efficiency of the Private vs. Public Storage”**

**Date:** 2012

**Source type:** Conference paper

Mazhelis, Fazekas, and Tyrväinen present a model for determining whether private storage or public cloud storage is the more cost effective option. Identifying a gap in the scholarship, their model includes the length of intervals between acquisition of private storage. They determine that public cloud storage is more cost effective when the intervals between assessments of private storage are longer. However, the authors also demonstrate that the acquisition interval at which private storage becomes more financially beneficial is also affected by other factors: the utility premium charged by the cloud provider, necessary storage redundancy, and the costs of transferring data to and from the cloud. The paper applies the model in an example using the storage needs of Oxford University Computing Services.

**Full citation:** Mazhelis, O., Fazekas, G., & Tyrvainen, P. (2012, June 24-29). Impact of storage acquisition intervals on the cost-efficiency of the private vs. public storage. Paper presented at 2012 IEEE 5th International Conference on Cloud Computing (CLOUD), Honolulu, HI. doi:10.1109/CLOUD.2012.101.

**Notes:** The paper references Walker, Brisken, and Romney and Khajeh- Hosseini et al. (see annotations) as well as Mastroeni and Naldi “Long-Range Evaluation of Risk in the Migration to Cloud Storage” (see notes for annotation Mastroeni and Naldi “Storage Buy-or-Lease Decisions in Cloud Computing Under Price Uncertainty”).

**Author: Naldi, M. & Mastroeni, L.**

**Title: “Cloud Storage Pricing: A Comparison of Current Practices”**

**Date:** 2013

**Source type:** Conference paper

The authors survey the current pricing of cloud storage by computing and comparing the unit price of different providers serving individual consumers as well as businesses. The providers surveyed are: Dropbox, SugarSync, IDrive, Google Drive, Carbonite, Symform, Mozy, and Amazon. The paper observes two types of pricing: tapering pricing (“declining block rate charge”) and bundling (or “quantity discount”) (p. 31). All the providers surveyed, except Amazon, use bundling. The authors determine the best models relative to each other through a two-part tariff approximation. For individual users, Google Drive and IDrive are the best options. For businesses, Amazon, Carbonite, and SugarSync appear to offer the best plans, though Dropbox is also a contender when factors beyond the calculation are considered.

**Full citation:** Naldi, M., & Mastroeni, L. (2013, April 20-24). Cloud storage pricing: a comparison of current practices. *Paper presented at 2013 International Workshop on Hot topics in Cloud Services, 4th ACM/SPEC International Conference on Performance Engineering*, Prague. doi:10.1145/2462307.2462315.

**Notes:** An earlier version of this paper includes CrashPlan in the survey, but it is shown to be one of the weaker competitors:

Mastroeni, L., & Naldi, M. (2012). Analysis of cloud storage prices. Retrieved from <http://arxiv.org/pdf/1207.6011v1.pdf>.

**Authors:** Naldi, M. & Mastroeni, L.

**Title:** “Economic Decision Criteria for the Migration to Cloud Storage”

**Date:** 2014

**Source type:** Journal article

The authors create a model for determining buy-or-lease decisions based Differential Net Present Value (DNPV). They also account for future changes, causing the DNPV to become a range of values. The model presents two decision criteria: the mean and median DPNV. To evaluate the effectiveness of these decision criteria, the authors apply three risk measures from the financial sector: probability of error, VaR, and CVaR. In the area of uncertainty, where DNPV is near zero, mean DPNV is shown to involve less error; however, the two decision criteria perform similarly otherwise. Applying the model, the article demonstrates that larger companies looking toward storage over the long term will benefit from a buy decision. Finally, the authors extend earlier work on pricing insurance options and find that a reasonably priced policy is possible for a small company.

**Full citation:** Naldi, M., & Mastroeni, L. (2014). Economic decision criteria for the migration to cloud storage. *European Journal of Information Systems*, 1-13. Retrieved from <http://www.palgrave-journals.com/ejis/journal/vaop/ncurrent/pdf/ejis201434a.pdf>.

**Notes:** The article references earlier work by Naldi and Mastroeni, Walker, Brisken, and Romney, Laatikainen, Mazhelis, and Tyrväinen, and Khajeh- Hosseini et al. (see annotations).

**Author:** Reichman, A.

**Title:** *File Storage Costs Less in the Cloud than In-House*

**Date:** 2011

**Source type:** Report

The report first considers possible uses of cloud storage and identifies file storage (i.e. “discrete packets of data and generally less performance critical,” p. 4) as the best candidate. It then explains how a quick comparison between buying storage and paying subscription fees for cloud storage does not produce accurate results. Instead, a large number of additional factors and fees need to be considered for either option. The report goes on give a generalised comparison between internal and cloud storage of 100 TB of data, demonstrating a 74 percent reduction in cost with the cloud. The spreadsheet tool used for these calculations is available through Forrester. However, the report tempers its assessment with cautionary advice about other issues involved in using the cloud.

**Full citation:** Reichman, A. (2011). *File storage costs less in the cloud than in-house*. Retrieved from

[http://media.amazonwebservices.com/Forrester\\_File\\_Storage\\_Costs\\_Less\\_In\\_The\\_Cloud.pdf](http://media.amazonwebservices.com/Forrester_File_Storage_Costs_Less_In_The_Cloud.pdf).

**Authors:** Rosenthal, D. C., Rosenthal, D. S., Miller, E. L., Adams, I. F., Storer, M. W., & Zadok, E.

**Title:** “Toward an Economic Model of Long-Term Storage”

**Date:** 2012

**Source type:** Conference poster description

Since the life of the data is longer than the media on which it is stored, storage decisions will be made at many points over the life of the data. Rosenthal et al. describe two models for understanding the cost of data storage. The first model examines the cost for a unit of hardware over time and calculates the total expenditures without accounting for the time value of money. Their second model looks to the concept of Discounted Cash Flow from economics, compensating for its limitations through Monte Carlo simulations. This model examines a unit of data transferring between storage media over time. The authors determine the amount of money that must be invested with the data at an interest rate which allows the investment to pay for the expenses incurred. They demonstrate the high probability that such an investment will be able to pay for storage for 100 years.

**Full citation:** Rosenthal, D. C., Rosenthal, D. S., Miller, E. L., Adams, I. F., Storer, M. W., & Zadok, E. (2012). Toward an economic model of long-term storage. Poster presented at FAST2012 Work-In-Progress. Retrieved from [http://static.usenix.org/events/fast/poster\\_descriptions/Rosenthaldescription.pdf](http://static.usenix.org/events/fast/poster_descriptions/Rosenthaldescription.pdf).

**Author:** Rosenthal, D. S., Rosenthal, D. C., Miller, E. L., Adams, I. F., Storer, M. W., & Zadok, E.

**Title:** “The Economics of Long-Term Digital Storage”

**Date:** 2012

**Source type:** Conference paper

Rosenthal et al. assert that, though there are many models for the cost of digital preservation, these models do not account for the possibility that future growth in storage capacity might not proceed according to Kryder’s Law. Ultimately, they seek to create a Monte Carlo model which accounts for the changes which might occur if storage growth no longer follows Kryder’s Law. The paper presents two prototype models, one following the life of a unit of hardware and the other following a unit of data as it migrates between media. It also compares the economic benefits of disk, tape, and solid state storage. Finally, the authors examine the current state of cloud storage. They observe that cloud pricing does not decrease as storage costs decrease (due to increased capacity following Kryder’s Law) and that the cost to transfer data is a barrier to changing service providers. The paper concludes that the cloud is not an appropriate option for long-term storage.

**Full citation:** Rosenthal, D. S., Rosenthal, D. C., Miller, E. L., Adams, I. F., Storer, M. W., & Zadok, E. (2012). The economics of long-term digital storage. Paper presented at

Memory of the World in the Digital Age, Vancouver, BC. Retrieved from <http://www.lockss.org/locksswp/wp-content/uploads/2012/09/unesco2012.pdf>.

**Authors: Rosenthal, D. S., & Vargas, D. L.**

**Title: "LOCKSS Boxes in the Cloud"**

**Date: 2012**

**Source type: Report**

The article responds to the proposition that libraries using LOCKSS boxes might benefit financially by moving to the cloud. The authors examine the technical architectures which could be configured for LOCKSS boxes using cloud services and determine the best option. They apply this architecture in an experiment using Amazon S3. However, because cloud storage pricing has not decreased according to the rate of increase in storage capacity (Kryder's Law), local disk storage is the cheaper option for long-term storage. The article proposes adjustments to Amazon S3, which would make using the service more cost effective. Their proposals are vindicated by the recent introduction of Amazon Glacier. Though Amazon has marketed Glacier to the digital preservation community, its access charges are problematic for integrity checking and, if Glacier pricing follows the example of S3, it will be a more expensive long-term option than local storage.

**Full citation:** Rosenthal, D. S., & Vargas, D. L. (2012). LOCKSS Boxes in the Cloud. Retrieved from: <http://www.lockss.org/locksswp/wp-content/uploads/2012/09/LC-final-2012.pdf>.

**Authors:** Walker, E., Brisken, W., & Romney, J.

**Title:** "To Lease or Not to Lease from Storage Clouds"

**Date:** 2010

**Source type:** Journal article

Walker, Brisken, and Romney identify a gap in the scholarship and apply preexisting buy-or-lease models for business to cloud storage. The authors based their work on established calculations for determining the Net Present Value (NPV) of purchased and leased assets. They present a model for resolving a buy-or-lease decision using NPV and accounting for "capital cost," "operating cost," "disc price trends," "disc replacement rates," and "disk salvage value" (pp. 46-47). They apply the model to three possible scenarios. In the case of a "single-user computer," (p. 47) the model demonstrates that, if storage is needed for less than four years, leasing is the more cost-effective option; however, purchasing is recommended for long-term storage. For medium-size companies, leasing cloud storage is the best option. Large enterprises (over a thousand servers) will find leasing storage to be cost-effective for up to nine years but purchasing storage becomes more economical for storage of longer duration.

**Full citation:** Walker, E., Brisken, W., & Romney, J. (2010). To lease or not to lease from storage clouds. *Computer*, 43(4), 44-50. doi: 10.1109/MC.2010.115.

**Notes:** This article extends Walker's earlier work in:

Walker, E. (2009). The real cost of a CPU hour. *Computer*, 42(4), 35-41. doi:10.1109/MC.2009.135.

**Authors:** Wang, J., Hua, R., Zhu, Y., Xie, C., Wang, P., & Gong, W.

**Title:** "C-IRR: An Adaptive Engine for Cloud Storage Provisioning Determined by Economic Models with Workload Burstiness Consideration"

**Date:** 2012

**Source Type:** Conference paper

The paper presents an approach that inputs projected future storage needs into a model using the Internal Rate of Return concept from economics to calculate whether or not it is financially beneficial to move to the cloud. The authors apply the model in an example using Amazon S3. The resulting data demonstrates that the financial benefits of storage in the cloud exceed buying new disk drives for small organisations but not for large organisations (enterprises with more than 1 TB or storage). Additionally, they propose a Burstiness Filter (referring to "cloud bursting" when a peak computing level is hit and data is transferred from the data center into the cloud)<sup>5</sup> to detect what work might be transferred to the cloud to increase cost savings.

**Full citation:** Wang, J., Hua, R., Zhu, Y., Xie, C., Wang, P., & Gong, W. (2012, June 28-30). C-

IRR: An adaptive engine for cloud storage provisioning determined by economic models with workload burstiness consideration. Paper presented at IEEE 7th International

---

<sup>5</sup> For the InterPARES Trust definition of "cloud bursting" see: [http://arstweb.clayton.edu/interlex/term\\_review.php?term=cloud](http://arstweb.clayton.edu/interlex/term_review.php?term=cloud).

Conference on Networking, Architecture and Storage (NAS), Xiamen, Fujian. doi:  
10.1109/NAS.2012.13.

**Notes:** The authors review the literature and cite Walker, Brisken, and Romney (see above annotation).



## Tier 2

Tier 2 includes sources which model or discuss the overall costs of cloud computing. Much of this research deals with storage as one of many cost considerations. Tier 2 also includes relevant studies on the costs of archival storage. However, literature on the costs of digital preservation was not covered exhaustively as InterPARES 3 General Study 16 already conducted research in this area (see below annotation). Also, more research on the costs of cloud computing exists in the field of Computer Science. This project only attempted to cover what appeared in our database searches (see appendix), but other research can be found through the literature reviews and reference lists in these sources.

**Author:** 451 Research

**Title:** *The Cloud Pricing Codex*

**Date:** 2013

**Source type:** Report

The report states that cloud service providers have fallen short of the imagined utility-style, usage-based pricing. 451 Research presents “a cloud pricing taxonomy” (p. 1) with eight categories of VM pricing. They observe that IaaS charges might also occur for a range of applications in the cloud, leading to unanticipated costs for the user. Further, the report finds that “only 64% of IaaS providers publish prices” (p.1).

**Full citation:** 451 Research. (2013). *The cloud pricing codex*. Retrieved from <https://451research.com/report-long?icid=2770>.

**Notes:** We did not have access to the full report. The summary is based on an executive overview available online. We found the report through a review in Business Source Complete:

Greengard, S. (2013). For enterprise IT, cloud pricing isn't simple. *CIO Insight*, 1.

**Author:** Ajeh, D.E., Ellman, J., & Keogh, S.

**Title:** “A Cost Modeling System for Cloud Computing”

**Date:** 2014

**Source type:** Conference paper

Ajeh, Ellman, and Keogh provide an overview of cloud computing, present the concept of cost modeling, and explain their own web-based tool: cloud computing modelling system (CCMS). CCMS is a tool to support decisions about cloud computing, operating through four phases. The tool begins with the “start-up phase” when the user inputs required information, followed by the “computation phase” in which the tool calculates the outputs given to the user in the “reporting phase” (p. 79). In the final “analytic phase,” the user assesses the results (p. 79). The authors proceed to demonstrate the use of the tool through three scenarios with hypothetical inputs. Through these tests they further find that particular usage patterns of computing resources translate to increased cost savings and that the longer the “economic life” of the infrastructure greater the financial benefit.

**Full citation:** Ajeh, D.E., Ellman, J., & Keogh, S., (2014, June 30- July 3). A cost modeling system for cloud computing. Paper presented at 14th International Conference on Computational Science and Its Applications (ICCSA), Guimaraes. doi: 10.1109/ICCSA.2014.24.

**Notes:** The authors reference the work of Khajeh- Hosseini et al. (see above tier 1 annotation).

**Author:** Allen, A., InterPARES 3 Project, Team Canada

**Title:** *General Study 16—Cost Benefit Models: Final Report*

**Date:** 2013

**Source type:** Report

The report presents a framework for considering the costs and benefits of digital preservation. It provides overviews of existing cost models for the range of activities involved in digital preservation, and many include the costs of storage. The report finds commonalities between the models and develops generalised frameworks for cost related activities and factors as well as the benefits of digital preservation. The authors point out that none of the models come from a strictly archival context, some dealing with library material or science data. The report concludes that further research is needed, since the models reviewed provide useful advice but do not provide the means by which to accurately predict cost.

**Full citation:** Allen, A., InterPARES 3 Project, Team Canada (2013). *General study 16—cost benefit models: final report*. Retrieved from

[http://www.interpares.org/display\\_file.cfm?doc=ip3\\_canada\\_gs16\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip3_canada_gs16_final_report.pdf).

**Notes:** General Study 16 also created an annotated bibliography of relevant research:

Kovynev, S., InterPARES 3 Project, Team Canada (2013). *General study 16—cost benefit models: annotated bibliography*. Retrieved from

[http://www.interpares.org/ip3/display\\_file.cfm?doc=ip3\\_canada\\_gs16\\_annotated\\_bibliography.pdf](http://www.interpares.org/ip3/display_file.cfm?doc=ip3_canada_gs16_annotated_bibliography.pdf).

**Author:** ARMA International

**Title:** *Guideline for Outsourcing Records Storage to the Cloud*

**Date:** 2010

**Source type:** Report

The ARMA report provides an overview of cloud computing and addresses issues surrounding cloud storage, including retention and disposition of data, legal concerns, and what to consider in evaluating vendors. The report states that long-term storage of records in the cloud is beyond its scope. However, it contains a section on cost, which details the cost benefits of cloud storage through its metered service, operational rather than capital expenditures, scalable and elastic qualities, and significant reduction of IT costs. The report also provides checklists and a questionnaire to assist records managers in evaluating cloud services.

**Full citation:** ARMA International (2010). *Guideline for outsourcing records storage to the cloud*. Retrieved from <http://www.arma.org/bookstore/files/ARMA-Outsourcing%20Cloud%20PDF-final.pdf>.

**Author:** Beargrie, N., Charlesworth, A., & Miller, P.

**Title:** *Guidance on Cloud Storage and Digital Preservation: How Cloud Storage can address the needs of public archives in the UK*

**Date:** 2014

**Source type:** Report

The report provides an overview of cloud computing, including security, legal issues, and cost, as well as advice for information professionals considering cloud options. It also contains summaries of cases studies on cloud storage conducted by the National Archives. The section devoted to cloud storage cost explains that, while cloud storage can be a cheaper option, it requires the archives to budget differently, as service providers typically bill on a monthly basis. However, this issue might be resolved through using specialised services, which have longer billing periods, or third party services for projecting cost over time. The report also includes a list of issues to consider and questions to ask about cost as part of an appended table.

**Full citation:** Beargrie, N., Charlesworth, A., & Miller, P. (2014). *Guidance on cloud storage and digital preservation: how cloud storage can address the needs of public archives in the UK*. Retrieved from:

<http://www.nationalarchives.gov.uk/documents/archives/cloud-storage-guidance.pdf>.

**Authors:** Biocic, B., Tomic, D., & Ogrizovic, D.

**Title:** “Economics of the Cloud Computing”

**Date:** 2011

**Source type:** Conference paper

The article provides general overviews of IaaS, PaaS, and SaaS as well as the criteria to consider in cloud adoption. The authors argue that it is not possible to create a generalised formula for calculating the various aspects of cloud adoption. If an organisation wishes to calculate costs for its own specific situation, it should assess private and public cloud adoption.

**Full citation:** Biocic, B., Tomic, D., & Ogrizovic, D. (2011, May 23-27). Economics of the cloud computing. Paper presented at *MIPRO 34th International Convention*, Opatija. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5967286](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5967286).

**Author:** Brumec, S., & Vrček, N.

**Title:** “Cost Effectiveness of Commercial Computing Clouds”

**Date:** 2013

**Source type:** Journal article

The authors present their calculating cloud computing cost effectiveness (CCCE) method. First, they demonstrate how to determine the resources needed in terms of the execution time of the application and the number of computers required. They then present formulas for calculating the total costs of the decision either to buy or to lease resources. The method is applied in three scenarios: an individual user, a small enterprise, and a large user. The scenarios demonstrate that leasing is the better option

for the individual user and small to medium enterprises; however, for large enterprises, purchasing in-house options is more cost effective. Finally, CCCE deals with the theory of trade relationships from economics, which assesses strong and weak buyers and vendors. The authors apply their method in a case study using the Croatian Information Systems of Higher Education Institutions.

**Full citation:** Brumec, S., & Vrčec, N. (2013). Cost effectiveness of commercial computing clouds. *Information Systems*, 38(4), 495-508. doi:10.1016/j.is.2012.11.002.

**Notes:** The article references Walker, Brisken, and Romney (see above tier 1 annotation).

**Author: California Digital Library**

**Title: “Total Cost of Preservation (TCP): Cost and Price Modeling for Sustainable Services”**

**Date:** 2013

**Source type:** White paper

The Total Cost of Preservation framework abstracts preservation activities into ten categories based on the Open Archival Information System model. Each of these categories has associated costs, and the report proceeds to detail two costing models. The “pay-as-you-go” (p. 6) model involves annual, per-unit costs incurred by the content owner and is most appropriate for institutions with steady, predictable funding. The “paid-up” (p. 7) pricing model is a one-time payment for all preservation activities and relies on predictions using discounted cash flow. This model is best used by organisations receiving funding for a limited time. The paper demonstrates mathematically calculating when one model is preferable to the other and offers suggestions for dealing with unexpected costs exceeding the predicted amount. Since the paper includes cloud service providers in an appendix discussing the real costs of long-term preservation, it appears that the model is applicable to cloud storage.

**Full citation:** California Digital Library (2013). *Total cost of preservation (TCP): cost and price modeling for sustainable services*. Retrieved from [https://wiki.ucop.edu/download/attachments/163610649/TCP-cost-price-modeling-for-sustainable-services-v2\\_1.pdf?version=4&modificationDate=1375721821000](https://wiki.ucop.edu/download/attachments/163610649/TCP-cost-price-modeling-for-sustainable-services-v2_1.pdf?version=4&modificationDate=1375721821000).

**Notes:** DSH Rosenthal et al. “The Economics of Long-Term Storage” is cited as a model similar to the “paid-up” model.

**Authors: Chang, Y. S., Lee, Y. K., Juang, T. Y., & Yen, J. S.**

**Title: “Cost Evaluation on Building and Operating Cloud Platform”**

**Date:** 2013

**Source type:** Journal article

The authors build on of the cost model developed by Walker, Brisken, and Romney (see tier one annotation) specifically for storage clouds and Tak, Urganonkar, and Sivasubramaniam’s more general cloud cost model (see below tier two annotation). They form their own model for comparing the cost of building an in-house data center and using a public cloud. The cost of the data center includes the cost of initially buying hardware, salaries for employees, maintenance of the hardware, power consumed by

the servers, and the cost of cooling the servers as well as the salvage value of the hardware. The price of a public cloud platform is modeled using the cost of Chunghwa Telecom's HiCloud and also includes the salary of an employee. The model is tested based on hypothetical enterprises of varying size. The authors find that a public cloud platform is beneficial for a one to two year time period but that private data centers benefit long-term usage.

**Full citation:** Chang, Y. S., Lee, Y. K., Juang, T. Y., & Yen, J. S. (2013). Cost evaluation on building and operating cloud platform. *International Journal of Grid and High Performance Computing (IJGHPC)*, 5(2), 43-53.

**Title: Cloudability**

**Date:** 2015

**Source type:** Website

Cloudability is a tool for monitoring and determining how to best manage cloud related finances. It offers a range of packages based on the amount of financial assets to be monitored.

**Full citation:** Cloudability (2015). Retrieved from: <https://cloudability.com/>.

**Title: Cloudyn**

**Date:** 2015

**Source type:** Website

Cloudyn is a SaaS product for planning, monitoring, and controlling cloud costs. It has deployments with several cloud service providers. Customers can select different versions ranging from the free "Lite" Cloudyn to a customised service.

**Full citation:** Cloudyn (2015). Retrieved from: <https://www.cloudyn.com/>.

**Author: Han, Y.**

**Title: "Cloud Computing: Case Studies and Total Costs of Ownership"**

**Date:** 2011

**Source type:** Journal article

Han first provides an overview of cloud computing, presenting a literature review of emerging scholarship on cloud computing in libraries as well as defining cloud computing, cloud services, and cloud providers. The article presents two case studies in which the author employed different cloud providers in library projects. In his subsequent analysis, Han breaks down the amounts involved in the total cost of ownership for both AWS and an in-house server for five years, which demonstrates significant cost savings with ASW. However, he also notes that digital libraries often need a very large amount of storage. Comparing TCO for 10TB in Amazon S3 versus physical storage onsite, the article shows in-house storage to be more cost effective than the cloud.

**Full citation:** Han, Y. (2011). Cloud computing: case studies and total costs of ownership. *Information Technology & Libraries*, 30(4), 198-206. Retrieved from <http://search.proquest.com.ezproxy.library.ubc.ca/docview/905719538?accountid=14656>.

**Author:** Haselmann, T., & Vossen, G.

**Title:** “EVACS: Economic Value Assessment of Cloud Sourcing by Small and Medium-sized Enterprises”

**Date:** 2014

**Source type:** Journal article

While cloud computing is generally considered financially beneficial for small and medium size enterprises (SMEs), the authors note that SMEs often encounter difficulties forming a business case for the cloud because they do not have the resources to adequately assess the issue. The authors present their Economic Value Assessment of Cloud Sourcing (EVACS) method, a three-step process for identifying viable projects for cloud services. In the first step, the enterprise assesses the project based on a set of indicators and contra-indicators. If the project is still a viable candidate, it is assessed by an additional set of “rules of thumb” (p. 3). The final step is an in-depth analysis by IT specialists. The authors list categories of costs to consider in this step: “Preparatory costs,” “Initialization costs,” “Operating costs,” and “Disinvestment costs.” In the case of existing products transferred to the cloud, they add four further categories: “Reengineering of existing systems,” “Migration of existing data,” “Integration of non-cloud systems,” and “Redundant operation of systems” (pp. 4-5).

**Full citation:** Haselmann, T., & Vossen, G. (2014). EVACS: economic value assessment of cloud sourcing by small and medium-sized enterprises. *Emisa Forum*, 34 (1), 18-31. Retrieved from [http://www.ngp.org.sg/cloud-asia2014/themes/tb\\_events\\_starter/sites/all/themes/tb\\_events/documents/EVACS-Economic-Value-Assessment-of-Cloud-Sourcing-by%20Small-and-Medium-sized-Enterprises.pdf](http://www.ngp.org.sg/cloud-asia2014/themes/tb_events_starter/sites/all/themes/tb_events/documents/EVACS-Economic-Value-Assessment-of-Cloud-Sourcing-by%20Small-and-Medium-sized-Enterprises.pdf).

**Notes:** The authors cite Walker, Brisken, and Romney (see above tier 1 annotation).

**Author:** Heinrich, H.

**Title:** “The PlanforCloud Calculator”

**Date:** 2013

**Source type:** Review

Heinrich favorably reviews the PlanforCloud calculator (5 out of 5) for use by libraries trying to understand the costs of using a cloud service. The tool was originally developed by Ali and Hassan Khajeh-Hosseini and purchased by RightScale. It models various aspects of a cloud deployment, including storage, and the specific usage scenario as well as the current prices of service providers.

**Full citation:** Heinrich, H. (2013). Tech services on the web: the PlanforCloud calculator <http://www.planforcloud.com/>. *Technical Services Quarterly*, 30(1), 120-121. doi: 10.1080/07317131.2013.735980.

**Notes:** The tool appears related to the work of Ali Khajeh-Hosseini on the Cloud Adoption Toolkit (see below annotation).

**Authors:** Hole, B., Lin, L., McCann, P., & Wheatley, P.

**Title: “LIFE<sup>3</sup>: A Predictive Costing Tool for Digital Collections”**

**Date:** 2010

**Source type:** Conference paper

The Life Cycle Collection Management project, which completed its third phase in 2010 (LIFE<sup>3</sup>), is a costing tool for library digitisation projects looking at the entire lifecycle of the material. Building on a model developed in an earlier phase of the project, LIFE<sup>3</sup> created a web-based tool for predicting cost, in which a user can enter information on a range of cost-related factors and receive a cost prediction. One of the many factors included in the tool is storage media, and cloud storage is listed as an option. The paper notes that the tool can be used to test the outcomes of using different forms of storage and storage vendors. The article goes on to describe other variables included in the tool as well as projections for future work.

**Full citation:** Hole, B., Lin, L., McCann, P., & Wheatley, P. (2010). LIFE<sup>3</sup>: a predictive costing tool for digital collections. Paper presented at iPRES 2010, 7<sup>th</sup> International Conference on Preservation of Digital Objects, Austria. Retrieved from <http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/hole-64.pdf>.

**Author: Instrumental, Inc. (for the Minnesota Historical Society)**

**Title: *Report on Digital Preservation and Cloud Services***

**Date:** 2013

**Source type:** Report

The report provides a break down of specific issues related to storing content in the cloud, followed by assessments of different service providers. These sections focus primarily on security and data integrity. Regarding storage cost, the report provides predictions based on current pricing models. However, because of inevitable changes to pricing models, it does not provide numbers but, instead, categorises costs as “low, medium, or high,” providing this assessment in a table with other information on cloud service providers. It also warns about the hidden costs in service provider contracts and lists some examples.

**Full citation:** Instrumental, Inc. (2013). *Report on Digital Preservation and Cloud Services*. Retrieved from [http://www.mnhs.org/preserve/records/docs\\_pdfs/Instrumental\\_MHSReportFinal\\_Public\\_v2.pdf](http://www.mnhs.org/preserve/records/docs_pdfs/Instrumental_MHSReportFinal_Public_v2.pdf).

**Notes:** The introduction references the work of David SH Rosenthal (see above tier 1 annotations), though it does not provide a specific citation.

**Author: ISACA**

**Title: “Calculating Cloud ROI: From the Customer Perspective”**

**Date:** 2012

**Source type:** White paper

The paper presents a framework for decision-makers to use in estimating the Return on Investment (ROI) of moving to the cloud. It first provides an overview of cloud computing and simple ROI calculations, presenting tables of the benefits, costs, and business challenges of cloud computing. The framework itself consists of three phases

each comprising a series of steps. The phases are: determining the costs and benefits of the cloud service, determining the current costs of services to be moved to the cloud, and using this information to calculate the ROI.

**Full citation:** ISACA. (2012). *Calculating cloud ROI: from the customer perspective*. Retrieved from <http://www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/Calculating-Cloud-ROI-From-the-Customer-Perspective.aspx>.

**Author:** Johnson, K., Reed, S., & Calinescu, R.

**Title:** “Specification and Quantitative Analysis of Probabilistic Cloud Deployment Patterns”

**Date:** 2012

**Source type:** Revised conference paper

Due to the scalable nature of cloud computing, cloud usage patterns are inevitably probabilistic. The authors present a model for determining cost and resource usage in the cloud which addresses this aspect of cloud computing. They use a probabilistic pattern modeling (PPM) approach which “formalizes cloud computing resources as probabilistic patterns and synthesizes Markov decision processes” (p. 158). The work is quantitatively verified using a probabilistic model checker (PRISM). The authors present a PPM tool that implements their approach as an open-source Java library. A test version of this tool is applied in a case study involving a hypothetical cloud customer, and the authors perform further experiments to test its scalability.

**Full citation:** Johnson, K., Reed, S., & Calinescu, R. (2012). Specification and quantitative analysis of probabilistic cloud deployment patterns. In *Hardware and Software: Verification and Testing* (pp. 145-159). Springer Berlin Heidelberg. doi:10.1007/978-3-642-34188-5\_14.

**Notes:** The authors reference Walker, Brisken, and Romney and Khajeh-Hosseini et al. (see above tier 1 annotations).

**Author:** Kratzke, N.

**Title:** “Cloud Computing Costs and Benefits: An IT Management Point of View”

**Date:** 2012

**Source type:** Book chapter

Kratzke first examines the benefits and shortcomings of cloud services based on the models COBIT, TOGAF, and ITIL. The author diagrams the positive, negative, or neutral impacts to the key components of each model. Further, though cloud providers are transparent in the billing of costs incurred, it is very difficult to predict costs prior to actually using the cloud service. Kratzke’s cost model uses the transparency in provider billing to predict costs for similar system architectures. Examining real-world billing reveals the following categories of usage costs: “Data Transfer,” “Data Storage,” “Processing,” “Requests,” and “Network” (p. 197-198). The paper further examines aspects of user interaction that affect cost. As the system architecture has a significant influence on cost, the model also includes a method for numerically calculating the similarity between system architectures.



**Full citation:** Kratzke, N. (2012). Cloud computing costs and benefits: an IT management point of view. In *Cloud Computing and Services Science* (pp. 185-203). Springer New York. doi:10.1007/978-1-4614-2326-3.

**Notes:** Kratzke cites Walker, Brisken, and Romney (see above tier 1 annotation).

**Authors:** Martens, B., & Teuteberg, F.

**Title:** “Decision-Making in Cloud Computing Environments: A Cost and Risk Based Approach”

**Date:** 2012

**Source type:** journal article

The article presents a comprehensive mathematical model to aid in the selection of cloud service providers by balancing cost and risk factors. The approach has theoretical foundations in transactional cost theory, resource-based theory, and relationship theory. The model incorporates a number of different cost factors: “IT service costs,” “Negotiation costs,” “Allocation costs,” “Coordination costs,” “Adoption costs,” “Maintenance costs,” and “Agency costs” (p. 881-883). The authors also consider three significant risks in cloud computing: “confidentiality, integrity and availability” (p. 883). The model is implemented using FICO Xpress Optimizer (a mathematical optimisation software tool) and applied in a simulation. This hypothetical scenario addresses the needs of a small/medium-size enterprise looking to use the cloud for storage. After examining three sourcing options, the authors conclude that the most cost effective cloud option might not be chosen when risk factors are also considered.

**Full citation:** Martens, B., & Teuteberg, F. (2012). Decision-making in cloud computing environments: A cost and risk based approach. *Information Systems Frontiers*, 14(4), 871-893. doi: 10.1007/s10796-011-9317-x.

**Author:** Miles, D.

**Title:** “Content in the Cloud- Making the Right Decision”

**Date:** 2012

**Source type:** White paper

The paper reports the results of a survey which asked organisations about a range of issues involved in cloud computing and concludes with specific recommendations. It notes that, along with IT and information professionals, financial departments are also hesitant about moving to the cloud. However, reducing cost is still seen as a primary drive for adopting cloud services. Major pro-cloud arguments include reduction of IT departments and the shift from capital expenditure to operational expenditure. Yet some CFOs are wary of the costs involved in renting cloud services for terms longer than five years. The survey also shows that IT is concerned that long-term cost of the cloud is not recognised.

**Full citation:** Miles, D. (2012). *Content in the cloud- making the right decision*.

Retrieved from <http://www.aiim.org/pdfdocuments/iw-content-in-the-cloud-2012.pdf>.

**Authors: Mohan Murthy, M. K., Sanjay, H.A., & Ashwini, J. P.**

**Title: “Pricing Models and Pricing Schemes of IaaS Providers: A Comparison Study”**

**Date:** 2012

**Source type:** conference paper

The paper explains different pricing models used by IaaS providers and compares specific providers. The authors identify two pricing models: the linear model (and its variations) where price increases linearly with the resources used and the “step model” in which cost decreases at certain amounts of resource usage. The authors further compare the pricing of cloud providers for storage and computational resources. Regarding storage, if the customer requires one TB or less of storage, RackSpace is the cheapest option; however, Amazon provides better pricing for storing more than one TB.

**Full citation:** Mohan Murthy, M. K., Sanjay, H. A., and Ashwini, J. P. (2012, August 3-5). Pricing models and pricing schemes of IaaS providers: a comparison study. Paper presented at *International Conference on Advances in Computing, Communications and Informatics*, Chennai, India. doi:10.1145/2345396.2345421.

**Author: Nanath, K., & Pillai, R.**

**Title: “A Model for Cost-Benefit Analysis of Cloud Computing”**

**Date:** 2013

**Source type:** Journal article

Nanath and Pillai present a three-level cost analysis framework. The first layer is a base cost estimate for the cloud service which considers nine cost components: “amortization, cost of servers, network cost, power cost, software cost, cooling cost, real estate cost, facility cost and support & maintenance cost” (p. 98). Layer two deals with the cost ramifications of an organisation’s data pattern, specifically analysing the time and demand constraints on computing resources. The third layer incorporates the specific cost considerations of the project under consideration. The article’s focus is limited to Amazon EC2. The model is applied in a survey of Indian IT firms; however, this study did not attempt to include layer three or the demand analysis of layer two. Overall, the results demonstrate that the cloud is financially beneficial for small and medium size organisations, but it is more cost effective for large enterprises to run their own data centers.

**Full citation:** Nanath, K., & Pillai, R. (2013). A model for cost-benefit analysis of cloud computing. *Journal Of International Technology & Information Management*, 22(3), 93-117. Retrieved from <http://search.proquest.com.ezproxy.library.ubc.ca/docview/1522799222?accountid=14656>.

**Authors:** Nielsen, A. B., Thirifays, A., & Kejser, U. B.

**Title:** “Costs of Archival Storage”

**Date:** 2012

**Source type:** Conference paper

The paper presents research conducted as part of the Cost Model for Digital Preservation (CMDP) conducted by the Danish National Archives, the Royal Library and the State and University Library. Using the OAIS model, CMDP breaks the cost of digital preservation into modules. The paper presents the Storage module. The module deals with activities critical to determining cost: receiving data, managing storage hierarchy, replacing media, error checking, disaster recovery, and providing access to data. It also accounts for the storage solution in terms of the device and media chosen. The authors find that storage cost depends on the volume of data as well as the organisation’s requirements for the storage of copies (i.e. if copies are stored in different geographic locations, on different devices, etc.).

**Full citation:** Nielsen, A. B., Thirifays, A., & Kejser, U. B. (2012, June 12-15). Costs of Archival Storage. In *Archiving 2012*. Paper presented at Archiving 2012: Preservation Strategies and Imaging Technologies for Cultural Heritage Institutions and Memory Organization, Copenhagen (pp. 205-210). Society for Imaging Science and Technology.

**Notes:** The paper cites D.S.H. Rosenthal’s model (see above tier 1 annotations) as presented on his blog (<http://blog.dshr.org/>).

**Author:** Rackspace Support

**Title:** “Clouconomics: The Economics of Cloud Computing”

**Date:** 2012

**Source type:** White paper

The paper details the cost benefits of cloud computing derived from four specific factors. It first includes “opportunity cost,” or the cost of not making a particular decision, in examining the decision to maintain a data center or move to the cloud. A cost benefit of cloud computing is that expenditures are operational instead of capital, and the paper details the specific advantages of operational over capital expenditures. The total cost of ownership in the cloud is also a benefit because it is possible to determine costs upfront as opposed to the unexpected costs of running a data center. Lastly, organisations benefit from the cloud because it allows them to focus on their core functions as opposed to investing effort in significant IT infrastructure.

**Full citation:** Rackspace Support (2012). *Clouconomics: the economics of cloud computing*. Retrieved from

[http://www.rackspace.com/knowledge\\_center/whitepaper/clouconomics-the-economics-of-cloud-computing](http://www.rackspace.com/knowledge_center/whitepaper/clouconomics-the-economics-of-cloud-computing).

**Author: Rutland, D.**  
**Title: “Cloud: Economics”**  
**Date: 2012**

**Source type:** White paper

Rutland discusses the general cost of cloud services and cautions against the assumption that using a public cloud is cheaper than dedicated hosting. It proceeds to discuss specific factors influencing cost. Firstly, if workloads fluctuate, there is a cost benefit in moving to the cloud. Regarding the widely propagated idea that the benefits of cloud services derive from the shift from capital to operational expenditure, the paper states that the issue is more complicated and that Total Cost of Ownership needs to be considered. It also includes the cost of migrating data to the cloud as well as the state of IT assets currently in use. Lastly, the paper demonstrates the importance of accounting for the time value of money in Return on Investment calculations and provides an example accounting for Net Present Value.

**Full citation:** Rutland, D. (2012). *Cloud: economics*. Retrieved from [http://www.rackspace.com/knowledge\\_center/sites/default/files/whitepaper\\_pdf/Cloud\\_Economics-%20Final%2006%2011%20123.pdf](http://www.rackspace.com/knowledge_center/sites/default/files/whitepaper_pdf/Cloud_Economics-%20Final%2006%2011%20123.pdf).

**Authors: Tak, B.C., Urganonkar, B., & Sivasubramaniam, A.**  
**Title: “To More or Not to Move: The Economics of Cloud Computing”**  
**Date: 2011**

**Source type:** Conference paper

The authors state that their study takes Walker, Brisken, and Romney’s work on Net Present Value calculations (see above annotation) a step further. First, they categorise cost components into “quantifiable” and “less quantifiable” as well as “direct” and “indirect” costs (p. 2). The study includes purely in-house and cloud options along with hybrid options for “vertical” (applications split into two subsets) and “horizontal” (applications replicated in the cloud) partitioning (p. 2). The study examines the affects on cost related to a number of factors: workload intensity, growth in hardware capacity (Moore’s Law), transferring data, storage, software licensing, and variation in demand. The authors main conclusions are: the cloud is a cost effective option for small organisations or organisations that are not currently growing; the cost of data transfer makes vertical partitioning expensive; and horizontal partitioning can be effective for dealing with peaks in resource demand.

**Full citation:** Tak, B.C., Urganonkar, B., & Sivasubramaniam, A. (2011). To move or not to move: the economics of cloud computing. Paper presented at *3rd USENIX conference on Hot topics in cloud computing*, Berkeley, CA, USA. Retrieved from [http://static.usenix.org/legacy/events/hotcloud11/tech/final\\_files/Tak.pdf](http://static.usenix.org/legacy/events/hotcloud11/tech/final_files/Tak.pdf) .

### Tier 3

Tier 3 consists of research relevant to understanding the costs of cloud computing. However, these sources deal with very specific situations (for example, science data) or present other aspects of cost modeling of limited relevance to this project.

**Authors:** Altmann, J., & Kashef, M. M.

**Title:** “Cost Model Based Service Placement in Federated Hybrid Clouds”

**Date:** 2014

**Source type:** Journal article

The paper notes the lack of sufficient scholarship on cost modeling for cloud computing. The authors conduct a systematic literature review and derive a set of 21 cost factors from this previous work. The paper then presents a comprehensive cost model for federated hybrid clouds, the most common cloud deployment according to the authors, and applies the model in a hypothetical scenario.

**Full citation:** Altmann, J., & Kashef, M. M. (2014). Cost model based service placement in federated hybrid clouds. *Future Generation Computer Systems*, 41, 79-90. doi:10.1016/j.future.2014.08.014.

**Notes:** The article cites Khajeh- Hosseini et al. (see above tier 1 annotation).

**Authors:** Buell, K. & Collofello, J.

**Title:** “Dynamic Cost Verification for Cloud Applications”

**Date:** 2012

**Source type:** conference paper

The paper is written for the purposes of software developers who need to know the costs of running their applications in the cloud. It presents static versus dynamic measurement of cost and determines that dynamic measurement is superior. Further, it presents different approaches for cost verification and concludes that the instrumentation approach, in which each part of the application has measurement layer for tracking cost, is the best option.

**Full citation:** Buell, K. & Collofello, J. (2012, July 15-20). Dynamic cost verification for cloud applications. Paper presented at *9th International Workshop on Dynamic Analysis, Minneapolis, MN*. doi:10.1145/2338966.2336802  
<http://doi.acm.org/10.1145/2338966.2336802>.

**Authors:** Deelman, E., Singh, G., Livny, M., Berriman, B., & Good, J.

**Title:** “The Cost of Doing Science on the Cloud: The Montage Example”

**Date:** 2008

**Source type:** Conference paper

The paper examines the cost benefits of running an astronomy application, Montage, on Amazon cloud services. The paper addresses three data management models: moving data to the cloud on-demand (and subsequently deleting it), storing data in the cloud where it can be shared, and regularly deleting unneeded files from storage. The

researchers run a set of simulations, testing different possible workflows. Regarding storage, the authors conclude that, for an application like Montage, storage cost is not significant in comparison to other costs in the cloud.

**Full citation:** Deelman, E., Singh, G., Livny, M., Berriman, B., & Good, J. (2008, November 15-21). The cost of doing science on the cloud: the Montage example. Paper presented at *2008 ACM/IEEE conference on Supercomputing*, Austin, TX. doi: 10.1109/SC.2008.5217932.

**Notes:** This study is referenced by Walker, Brisken, and Romney (see above tier 1 annotation).

**Author:** Franceschelli, D., Ardagna, D., Ciavotta, M., & Di Nitto, E.

**Title:** “SPACE4CLOUD: A Tool for System PerformAnce and Cost Evaluation of CLOUD Systems”

**Date:** 2013

**Source type:** Conference paper

The authors note the difficulty in determining which cloud provider fits the customer’s financial needs. They state that the current tools for predicting cost and performance for the purposes of software developers are not suited for modeling a cloud environment. They present research in the area of Model-Driven Quality Prediction, extending the Palladio Framework and mapping cloud meta-models to this framework. From this work, the authors derive their tool, System PerformAnce and Cost Evaluation for Cloud (SPACE4CLOUD) and apply it in a case study using the Banking suite of SPECweb2005. The model is shown to be successful for modeling light workloads but conservative for heavy workloads.

**Full citation:** Franceschelli, D., Ardagna, D., Ciavotta, M., & Di Nitto, E. (2013, April 21-24). SPACE4CLOUD: a tool for system performance and cost evaluation of cloud systems. Paper presented at *2013 International Workshop on Multi-Cloud Applications and Federated Clouds, Prague*. doi:10.1145/2462326.2462333.

**Authors:** Mazhelis, O., & Tyrväinen, P.

**Title:** “Economic Aspects of Hybrid Cloud Infrastructure: User Organization Perspective”

**Date:** 2012

**Source type:** Journal article

Mazhelis and Tyrväinen present a model for examining the cost effectiveness of hybrid clouds. For fixed unit prices, a hybrid cloud is superior to public or private cloud options. The authors also consider the costs of data communication. If more data is transferred, more computing capacity should be distributed to the private cloud component of the hybrid cloud. However, if unit prices involve a quantity discount (i.e. the price is discounted for a large purchase of a certain resource) either fully public or fully private cloud options are preferable. The in-house option is superior if there is a steady demand (as opposed to peaks in demand) for resources. These conclusions are demonstrated through numerical experiments. The authors note exploration of storage costs as an area for further research.

**Full citation:** Mazhelis, O., & Tyrväinen, P. (2012). Economic aspects of hybrid cloud infrastructure: user organization perspective. *Information Systems Frontiers*, 14(4), 845-869. doi: 10.1007/s10796-011-9326-9.

**Notes:** This article is cited in Laatikainen, Mazhelis, and Tyrväinen and Mazhelis, Fazekas, and Tyrväinen (see above tier 1 annotations).

**Authors:** Mian, R., Martin, P., Zulkernine, F., & Vazquez-Poletti, J.L.

**Title:** “Estimating Resource Costs of Data-Intensive Workloads in Public Clouds”

**Date:** 2012

**Source type:** Conference paper

The paper presents a costing model for a public cloud with pay-as-you-go pricing. The model consists of compute, storage, network (accessing stored resources), and penalty costs. The model is tested hypothetically in an instance of Amazon EC2. Three experiments are conducted considering three variables, each experiment varying one variable while holding the others constant. The three variables are: workloads, VM types, and SLOs enforced. In all the experiments, storage costs are found to be low.

**Full citation:** Mian, R., Martin, P., Zulkernine, F., & Vazquez-Poletti, J.L. (2012, December 3-7). Estimating resource costs of data-intensive workloads in public clouds. Paper presented at *10th International Workshop on Middleware for Grids, Clouds and e-Science, Montreal, Quebec*. doi:10.1145/2405136.2405139.

**Authors:** Ruiz-Alvarez, A. & Humphrey, M.

**Title:** “A Model and Decision Procedure for Data Storage in Cloud Computing”

**Date:** 2012

**Source type:** Conference paper

The paper presents a model for use in decision-making for data allocation and computation services (e.g. data analysis applications) in the cloud. A range of users' needs and the storage capabilities of cloud providers are input into the model, which accounts for cost, access latency, and bandwidth. The authors then present software (an integer linear programming - ILP - solver) which calculates solutions, and demonstrate the scalability of their model. They implement the tool in examples using the scientific applications BLAST (Basic Local Alignment Search Tools, a bioinformatics algorithm used for genetic sequence searching) and MODIS (an Azure application for processing satellite data). For BLAST, the authors adapt their model to include budget constraints. In the MODIS case, they include average computational length in order to demonstrate the fastest completion of jobs on a given budget.

**Full citation:** Ruiz-Alvarez, A. & Humphrey, M. (2012, May 13-16). A model and decision procedure for data storage in cloud computing. Paper presented at *12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, Ottawa, ON. doi:10.1109/CCGrid.2012.100.

**Notes:** For an earlier version of the system see:

Ruiz-Alvarez, A. & Humphrey, M. (2011, June 8-10). An automated approach to cloud storage service selection. Paper presented at *2nd International Workshop on Scientific Cloud Computing*, San Jose, California. doi:10.1145/1996109.1996117.

**Authors:** Walterbusch, M., Martens, B., & Teuteberg, F.

**Title:** “Evaluating Cloud Computing Services from a Total Cost of Ownership Perspective”

**Date:** 2013

**Source type:** Journal article

The authors present a model for calculating the total cost of ownership (TCO) for start-up companies using services in a public cloud. The model follows the “life cycle” of the cloud service, comprising the phases: “initiation” (the decision to move to the cloud), “evaluation” (assessing cloud providers), “transition,” “operation,” and “dissolution” (p. 618). The article presents formulas for determining a series of cost factors. Activities that incur cost are: making the decision to move to the cloud; selecting the service provider; service charges associated with IaaS, PaaS, and SaaS respectively; transferring data and setting up the service; support services; training; maintaining and modifying the service; possible system failure; and transferring data out of the cloud. Determining TCO is the sum of these factors. The model is assessed through a hypothetical example of a start-up company using an IaaS service and through interviews with experts. It is also applied in a software tool publically accessible online.

**Full citation:** Walterbusch, M., Martens, B., & Teuteberg, F. (2013). Evaluating cloud computing services from a total cost of ownership perspective. *Management Research Review*, 36(6), 613-638. doi:10.1108/01409171311325769.

**Notes:** The tool is still available online and the website is up-to-date (<http://www.cloudservicemarket.info/tools/tco.aspx>).