ECONOMIC AND TECHNICAL ASSESSMENT OF CLOUD COMPUTING

Serdar Yarlikas[1] and Semih Bilgen[2]
[1]Informatics Institute, Department of Information Systems, Middle East Technical University, 06800, Ankara, Turkey
[2]Department of Electrical and Electronics Engineering, Middle East Technical University, 06800, Ankara, Turkey

Keywords: Cloud computing, Economic variables, Technical variables.

Abstract: This position paper presents the current state of the research that aims to assess the economic and technical benefits of cloud computing and to uncover best practices and lessons learned by service providers as well as clients. In the first phase of the study, based on a review of the recent literature on cloud computing principles, applications, achievements and challenges, variables that determine cloud performance and benefits have been identified. Then, axioms and hypotheses on the interrelationships between the variables have been formulated. Verification of the conceptual network of axioms and hypotheses entails exploratory and validatory case studies. The paper elucidates the strengths and weaknesses, while pointing out the opportunities and threats involved in cloud operation and adaptation, specifically in private organizations.

1 INTRODUCTION

We examine cloud computing practices from technical and economic viewpoints. We construct a conceptual framework of relationships between the variables related to enablers and obstacles of migration to cloud environments. Once the hypotheses on those relationships are formulated based on a survey of the relevant literature, they will be further investigated through case studies.

2 THEORETICAL FRAMEWORK

Within the context of this study, cloud computing is defined as a mode of Internet-based large-scale distributed computing whereby shared software and information resources are provided to users over dynamically scalable and virtualized computing facilities. In this mode of computing, applications can be accessed over the Internet without prior installation, as Software as a Service (SaaS).

Different web-based enterprise solutions such as Customer Relations Management (CRM), Enterprise Resource Planning (ERP), Human Resources Management (HRM), Business Process Management (BPM), and Expense Management (EM) will also be investigated if they have been adapted to cloud platforms, and as such, can be classified as SaaS solutions. Within the literature that discusses the critical deterrents for cloud computing, some authors have concentrated on economic challenges, whereas others have also considered technology related issues (Armbrust, et al., 2010; Leavitt, 2009; Jiang and Yang, 2010).

2.1 Technical Variables

The technical variables to be studied have been identified as below. Unless specified otherwise, possibly in arithmetical terms, these are generally to be measured in a categorical scale such as (low, medium, high). When categorized otherwise, an ordering is defined on the categories so that the hypotheses that refer to such ordering can be validly presented as in Section 3.

Security refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing. Cloud security (CS) also includes data protection, identity management, physical and personnel security, application security and privacy (Wang, 2009).

Network virtualization (NV) is a method of combining the available resources in a network by splitting up the available bandwidth into channels, each of which is independent from the others, and each of which can be assigned (or reassigned) to a particular server or device in real time. Possible

Yarlikas S. and Bilgen S..
ECONOMIC AND TECHNICAL ASSESSMENT OF CLOUD COMPUTING.
DOI: 10.5220/0003960701670172
In *Proceedings of the 2nd International Conference on Cloud Computing and Services Science* (CLOSER-2012), pages 167-172
ISBN: 978-989-8565-05-1

levels of virtualization have been described, in increasing levels, as para-virtualization, hardware-assisted virtualization, live-migration, and pause-resume. (Uhlig, et al., 2005; Clark, et al., 2005; Travostino, et al., 2006; Padala, 2010).

Cloud deployment type (CDT) refers to the ownership and operational categorization of the computing infrastructure. Four major deployment types are: public, private, community and hybrid clouds (Chebrolu, 2011).

Upgradeability refers to the capability of modifying network structure to improve the performance in an easy and efficient way (Gutierrez, Riaz, Pedersen, Madsen, 2009).

*Manageability (*MNG*)* is defined as the collective processes of deployment, configuration, optimization, and administration during the lifecycle of cloud computing systems and services. Related metrics are proposed as checklist of manageability functions, number of steps to manage towards desired state, time to manage, documentability, elasticity of management, availability and continuity of management and ease of use (Cook, et al., 2011).

Alternatively, manageability means the ability to have visibility and control over services and usage, enabled by cloud programmatic management interfaces, cloud web management capabilities, self service provisioning (Citrix, 2010).

Availability is the percentage of time the cloud computing system is available for use via possibly multiple cloud computing providers (Armbrust et al., 2010; Helft, 2009; Beard, 2009).

Scalability (SCL) refers to the ability of a system to grow in one or more dimensions as more resources are added to the system. These dimensions include the number of concurrent users that can be supported and the number of transactions that can be processed in a given unit of time (WebLogic, 2011; SearchCloudComputing, 2011; CCD, 2011).

Agility (AGL) refers to the ability to rapidly change the cloud application to meet user requirements. (Citrix, 2010).

Operational efficiency (OEF) is the capability of a cloud vendor to deliver cloud services to its customers in the most cost-effective manner possible while still ensuring high quality of its service and support (Webopedia, 2011).

Data locality (DL) refers to the phenomenon that the collection of the data locations referenced in a short period of time in a running computer often consists of relatively well predictable clusters (Denning, 2005).

Resilience (fault tolerance) (RSL) is the property that enables a system to continue operating properly in the event of the failure of some of its components (SunGard, 2010).

Isolation failure refers to the failure of mechanisms that separate data that belong to different organizations sharing a cloud (Harris and Alter, 2010).

Concurrency (CCRCY) is a property of systems in which several computations are executing simultaneously, and potentially interacting with each other (Expert Group, 2010; Kovachev et al., 2011).

Vendor lock in degree (VLID) denotes the level of the customer dependency on a vendor for products and services (Armbrust et al., 2010; The Open Group, 2010).

Cloud performance (CP) is determined by various metrics referring to network performance, application performance and data center performance. These three performance components, are, in turn, influenced by all software and hardware infrastructures. Network performance is measured mainly by throughput and latency metrics. (Li, et al., 2011; Myerson, 2011).

Cloud infrastructure performance, on the other hand, can be evaluated with such metrics as CPU (Central Processing Unit) utilization and application traffic (Henkel, et al., 2007

Distance (DIS) is the relative location of the cloud customer to the cloud providers and the other cloud customers. Distance impacts response times, thus cloud application performance (Weinman, 2011).

2.2 Economical Variables

The economic variables to be studied have been identified as follows:

CAPEX (Capital Expenditure) includes the cost of entire data center infrastructure, including servers, storage arrays, software licenses (when needed), routers, and load-balancers. *CAPEX* means funds used by a company to acquire or upgrade physical assets such as equipment. (The Open Group, 2010).

Total cost of ownership (TCO) can be defined arithmetically as (Neville, 2010; The Open Group, 2010):

TCO= (Network management expenses+ the expenses of services)/ (number of customers of service providers), or

TCO= (Network costs+ Router/switch costs+personnel costs + training costs + hardware costs + software costs + endpoint costs+ energy costs+ facilities costs)/ (number of customers of service providers).

Outage duration cost (ODC) is the product of the length of outage duration, the number of customers served, and a fixed customer outage duration cost (Kuntz, et al., 2001).

Total cost of outages= f (duration of outage, size of data center, business disruption, lost revenues, IT equipment failure, accidental/human error), where

Duration of outage= f (length of outage duration, the number of customers served, fixed customer outage duration cost). Outage duration cost refers to the cost of losing access to the data, accrued when the system fails to provide or perform its primary function in a period of time. Outage duration cost can also be calculated by multiplying outage duration hours with amount of money lost per hour (Newvem, 2011; Kuntz et al., 2001).

Switching cost refers to the effort and expense involved in switching from one cloud provider to another cloud provider in terms of cloud services and products (Klemperer, 1987).

Cloud migration cost is the cost of carrying all applications from the current computing environment to cloud environment. It can be defined mathematically as follows (The Open Group, 2010; Blaisdell, 2011):

Cloud migration cost=Setup costs + Migration costs + Ongoing costs

Setup costs+ migration costs= The costs of moving application and business moving to the cloud.

Ongoing costs= f (CPU time, GB of RAM, terabyte of storage)

Data replication cost refers to the cost of sharing data in order to ensure consistency between redundant resources, such as hardware and software components, to improve reliability, fault tolerance, accessibility when the data is stored on multiple storage devices (CNTC, 2003). Of the many factors that have a bearing on data replication costs, support for heterogeneous storage, licensing fees, and network bandwidth requirements are the most significant (CNTC , 2003).

3 AXIOMS AND HYPOTHESES

In this section, relations between variables that are not specific to cloud systems or that are self-evident are listed as axioms, to be followed by cloud-specific relations referred as hypotheses to be verified by case studies.

3.1 Axioms

A1: In case of failure of mechanisms that separate data of organizations sharing a cloud, security decreases (Harris and Alter, 2010).

A2: Vendor lock-in degree causes switching cost increase. (Armbrust et al., 2010; The Open Group, 2010).

A3: If availability increases, outage duration cost decreases because the system can perform its primary function and the effectiveness of cloud applications also increases (The Open Group, 2010; Wilson, 2008).

A4: Increasing agility reduces cloud application migration cost (Khalidi, 2011).

A5: Increasing resilience increases availability (SunGard, 2010).

A6: If scalability increases, cloud performance increases (CCD, 2011).

A7: If availability increases, scalability increases (Armbrust et al., 2010; Helft, 2009; Beard, 2009; CCD, 2011; Varia, 2010).

A8: Agility increases upgradeability (Kundra, 2011).

3.2 Hypotheses

Hereafter, a (-) or (+) between two variables denote that an increase in the first variable causes a decrease or increase, respectively, in the second one.

H1a: Public cloud deployment (PCD) (-) vendor lock-in degree

The degree of vendor lock-in depends on the type of cloud deployment (Chebrolu, 2011), i.e. private, community, public or hybrid cloud (Chebrolu, 2011). Public cloud deployment causes lower vendor lock-in degree because cloud vendor provides applications available to general public in this deployment type and so cloud customer dependency on cloud vendor decreases.

H1b: Private cloud deployment (PRCD) (+) vendor lock-in degree

Private cloud deployment increases vendor lock-in degree (Chebrolu, 2011).

H2: Network virtualization (-) security

In the cloud environment, all resources are not virtualized. If a user can access the data of other users, this is considered “problematic virtualization” hence, network virtualization causes reduction of security (Armbrust et al., 2010).

H3: Level of virtualization (LOV) (+) manageability

Virtualization generally enables each application to be encapsulated such that they can be configured, deployed, started, migrated, suspended, resumed, stopped, and thus provides better manageability (Giordanelli and Mastroianni, 2010).

H4: Manageability (-) outage duration cost

In case of manageability problems, the system can not work until the fault is found, and so outage duration cost increases (The Open Group, 2010).

H5: Data locality (+) operational efficiency

Data locality minimizes the amount of data movement which increases efficiency (Giordanelli and Mastroianni, 2010).

H6: Data locality (+) cloud performance

Data locality improves end-application performance which increases cloud performance (Giordanelli and Mastroianni, 2010).

H7: Data locality (+) scalability

(Giordanelli and Mastroianni, 2010).

H8: Resilience (fault tolerance) (+) cloud performance

(SunGard, 2010).

H9: Scalability (+) operational efficiency

(CCD, 2011).

H10: Scalability (+) agility

Scalability, especially when dynamic, increases agility (Citrix, 2010).

H11: Distance (-) cloud performance

(Weinman, 2011).

H12: Concurrency (-) security

(Expert Group, 2010; Kovachev et al., 2011). (Anderson, 2008).

The conceptual network of the hypotheses is presented in Figure 1:

4 CONCLUSIONS

Based on the literature review, strengths achieved via cloud operation and adaptation have been identified as higher resource utilization, efficient resource management, increased agility, reduction in capital expenditure and outage duration cost, increased availability, decrease on data replication cost, ease of access, information sharing, improved security. On the other hand, weaknesses such as authentication, ineffective and problematic network virtualization, vendor-lock in, increase on switching cost may also result. Threats are identified as isolation failure, data protection, loss of governance, difficulty in providing innovative processes, whereas opportunities such as virtualization and scaling structures of cloud and resilience can arise.

Obstacles of cloud computing are scaling oriented issues, data security related issues, data lock-in and possible loss of system and applications availability.

Enablers of cloud computing have been identified as virtualization, network services, high speed networking and applications and open source structure.

At the current stage of this research, 8 axioms and 12 hypotheses have been formulated constituting a conceptual network of relationships. Verification of the hypotheses will be based on exploratory and validatory case studies. Matured hypotheses will then be studied with cloud service providers as well as clients in validatory case studies.

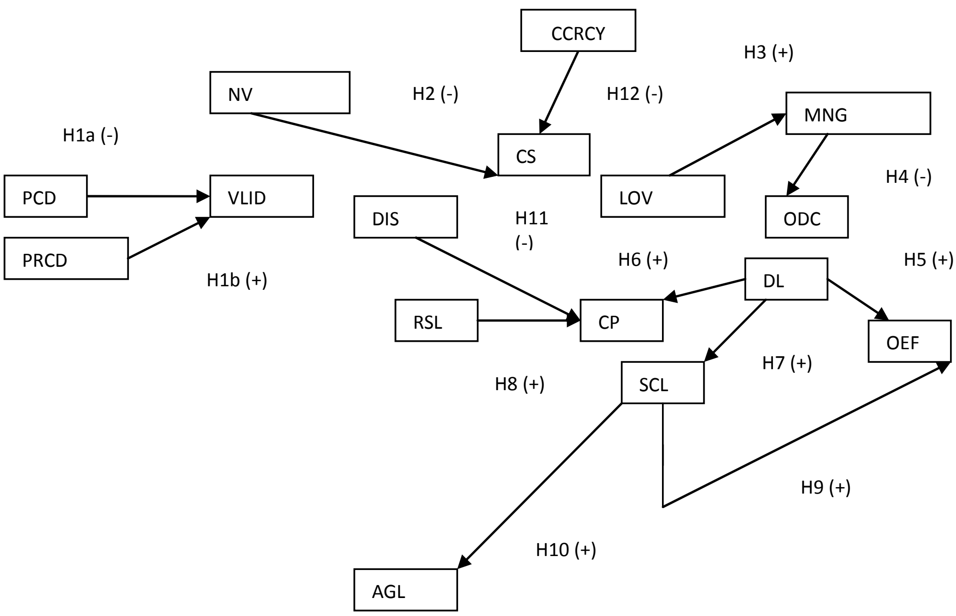

Figure 1: Conceptual Network of Hypotheses.

REFERENCES

Anderson, R. (2008). *Security Engineering: A Guide to building dependable distributed systems* (2nd ed.). Wiley.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D, Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50-58.

Beard, H. (2009). *Cloud Computing Best Practices for managing and measuring processes for on-demand computing* (ITIL). London, UK.

Blaisdell, R. (2011). *How to determine how much your cloud costs*. Retrieved December 28, 2011, from, http: //www.rickscloud.com/how-to-determine-how-much-your-cloud-costs/

Chebrolu, S. B. (2011). Assessing the Relationships among Cloud Adoption, Strategic Alignment and IT Effectiveness. *Journal of Information Technology Management*, 22(2), 13-29.

Citrix (2010). *Is your load balancer cloud ready*. Retrieved December 28, 2011, from, http://images01.insight.com/media/pdf/0311Whitepaper2NetScalerIsYourLoadBalancerCloudReady.pdf

Clark, C., Fraser, K., Hand, S., Hansen, J. G., Jul, E., Limpach, C., Pratt, I. and Warfield, A. (2005). Live migration of virtual machines. *Proceedings of the 2nd ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 273–286.

Cloud Computing Development (CCD, 2011). *The Secret to Doing Cloud Scalability Right.* Retrieved December 28, 2011, from, http://www.cloudcomputingdevelopment.net/the-secret-to-doing-cloud-scalability-right/

Computer Network Technology Corporation (CNTC, 2003). *Data Replication Considerations.* Retrieved December 28, 2011, from http://www.parallelprose.com/parallelprose/white_papers/CNT%20Replication..pdf

Cook, N., Milojicic, D., Talwar, V. (2011). *Cloud Management.* Retrieved December 28, 2011, from http://www.hpl.hp.com/techreports/2011/HPL-2011-238.pdf

Denning, P. J. (2005). The Locality Principle. *Communications of the ACM*, 48(7), 19–24.

Expert Group (2010). *The Future of Cloud Computing. Opportunities for European Cloud Computing Beyond 2010*. Retrieved August 21, 2011, from http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-report-final.pdf

Giordanelli, R., Mastroianni, C. (2010). *The Cloud Computing Paradigm: Characteristics, Opportunities and Research Issues*. Retrieved August 12, 2011, from http://grid.deis.unical.it/papers/pdf/RT-01-2010-CloudSurvey.pdf

Gutierrez, J. M., Riaz, T, Pedersen, J. M. and Madsen, O. B. (2009). Upgradeability and predictability analysis for mesh topologies in optical distribution networks. *Proceedings of the Sixth international conference on Wireless and Optical Communications Networks*, 369-373.

Harris, J. G., Alter, A. E. (2010). *Six Questions Every Executive Should Ask about Cloud Computing*. Retrieved August 13, 2011, from http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture_6_Questions_Executives_Should_Ask_about_Cloud_Computing.pdf

Helft, M. (2009). *Google confirms problems with reaching its services*. Retrieved July 24, 2011, from http://bits.blogs.nytimes.com/2009/05/14/google-confirms-problems-accessing-its-services/

Henkel, J., Parameswaran, C., Cheung, N. (2007). *"Application Specific Embedded Processors" in "Designing Embedded Processors"*. Springer, 3-23.

Jiang, J., Yang, G. (2010). *"Examining Cloud Computing from the Perspective of Grid and Computer-Supported Cooperative Work" in "Cloud Computing Principles, Systems and Applications"*. London: Springer, 63-76.

Khalidi, Y. A. (2011). Building a Cloud Computing Platform for New Possibilities. *IEEE Computer Society*, 44(3), 29-34.

Klemperer, P. D. (1987). Markets with Consumer Switching Costs. *Quarterly Journal of Economics*, 102, 375-394.

Kovachev, D., Cao, Y., Klamma, R. (2011). *Mobile Cloud Computing: A Comparison of Application Models*. Retrieved August 23, 2011, from http://arxiv.org/ftp/arxiv/papers/1107/1107.4940.pdf

Kundra, V. (2011). *Federal Cloud Computing Strategy*. Retrieved August 18, 2011, from http://www.cio.gov/documents/federal-cloud-computing-strategy.pdf

Kuntz, P. A, Christie, R. D, Venkata, S. S. (2001). A Reliability Centered Optimal Visual Inspection Model for Distribution Feeders. *IEEE Transactions on Power Delivery*, 16(4), 718-723.

Leavitt, N. (2009). Is Cloud Computing Really Ready for Prime Time? *IEEE Computing Magazine*, 42(1), 15-20.

Li, A., Yang, X., Kandula, S., Zhang, M. (2011). Comparing Public-Cloud Providers. *IEEE Computer Society*, 15(2), 50-53.

Myerson, J. M. (2011). *Craft a cloud performance metrics policy*. Retrieved December 28, 2011, from, http://www.ibm.com/developerworks/cloud/library/cl-cloudperformmetrics/

Neville, J. (2010). *INTUG Cloud Computing Position Paper.* Retrieved December 28, 2011, from, http://intug.org/wp-content/uploads/2009/03/Position-Paper-Cloud-Computing.pdf

Newvem (2011). *Availability: Story of the Inevitable Outage of the Cloud*. Retrieved December 28, 2011, from, http://blog.newvem.com/main/2011/08/availability-story-of-the-inevitable-outage-of-the-cloud.html

Padala, P. (2010). *Understanding Live Migration of Virtual Machines.* Retrieved December 1, 2011, from http://ppadala.net/blog/2010/06/understanding-live-migration-of-virtual-machines/

SearchCloudComputing (2011). *Cloud Computing from A to Z.* Retrieved December 28, 2011, from, http://searchcloudcomputing.techtarget.com/feature/Cloud-computing-from-A-to-Z

SunGard Availability Services UK Ltd. (2010). *Resilience is the Key to Success*. Retrieved August 24, 2011, from http://www.sungard.co.uk/Documents/Cloud-Computing-White-Paper-SunGard.pdf

The Open Group. (2010). *Building Return on Investment from Cloud Computing: Cloud Computing Key Performance Indicators and Metrics*. Retrieved March 19, 2011, from http://www.opengroup.org/cloud/white papers/ccroi/kpis.htm

Travostino, F., Daspit, P., Gommans, L., Jog, C., Laat, C., Mambretti, J., Monga, I., van Oudenaarde, B., Raghunath, S., and Wang, P. Y. (2006). Seamless live migration of virtual machines over the man/wan. *Future Generation Computer Systems*, 22(8), 901–907.

Uhlig, R., Neiger, G., Rodgers, D., Santoni, A. L., Martins, F. C. M., Anderson, A. V., Bennett, S. M., Kagi, A., Leung, F. H., Smith, L. (2005). Intel Virtualization Technology. *IEEE Computer Society*, 38(5), 48-56.

Varia, J. (2010). *Amazon Web Services- Architecting for the Cloud: Best Practices*. Retrieved August 15, 2011, from http://media.amazonwebservices.com/AWS_Cloud_Best_Practices.pdf

Wang, C. (2009). *Cloud Security Front and Center*. Retrieved March 1, 2011, from http://blogs.forrester.com/security_and_risk/2009/11/cloud-security-front-and-center.html

WebLogic (2011). *WebLogic Server Performance Tuning Guide*. Retrieved December 28, 2011, from, http://docs.oracle.com/cd/E13222_01/wls/docs51/admindocs/tuning.html

Webopedia (2011). *Operational Efficiency*. Retrieved December 28, 2011, from http://www.webopedia.com/TERM/O/operational_efficiency.html

Weinman, J. (2011). *Axiomatic Cloud Theory*. Retrieved December 28, 2011, from, http://www.joeweinman.com/Resources/Joe_Weinman_Axiomatic_Cloud_Theory.pdf

Wilson, S. (2008). *AppEngine outage*. Retrieved March 12, 2011, from http://www.cio-weblog.com/50226711/appengine_outage.php