8-2010

Privacy and Security Practices in the Arena of Cloud Computing - A Research in Progress

Srilakshmi Ramireddy
University at Buffalo, sr229@buffalo.edu

Rajarshi Chakraborthy
University at Buffalo, rc53@buffalo.edu

T.S. Raghu
Arizona State University, raghu.santanam@asu.edu

H. Raghav Rao
University at Buffalo, mgmtrao@buffalo.edu

Privacy and Security Practices in the Arena of Cloud Computing - A Research in Progress

Srilakshmi Ramireddy
University at Buffalo
sr229@buffalo.edu

Rajarshi Chakraborthy
University at Buffalo
rc53@buffalo.edu

T. S. Raghu
Arizona State University
Raghu.Santanam@asu.edu

H. Raghav Rao
University at Buffalo
mgmtrao@buffalo.edu

ABSTRACT

Cloud Computing has become an important player in the field of IT infrastructure outsourcing. The convenience this new computation model brings can however be easily balanced out by the uncertainty behind the "cloud". When so many organizations in the private and public sector are looking at cutting costs and resorting to Cloud services, it becomes imperative to investigate the amount of assurance the Cloud vendors give to their clients. In this paper we have thus explored the assurance related information available on the websites of some contemporary Cloud vendors and have tried to investigate if the characteristics of a vendor play any significant role in the vendor's decision to adopt certain information assurance practices, specifically in the context of Privacy and Security.

Keywords

Cloud Computing, Information Assurance, Cloud Vendors, ANOVA, emphasis index, assurance unavailability index.

INTRODUCTION

Cloud Computing approach offers dynamically scalable and often virtualized resources as services over the Internet (Behnia, 2009). Numerous conceptualizations of Cloud Computing alternatively emphasize "utility Computing," "services oriented Computing," or "on-demand Computing" as defining features of Cloud Computing approach (Mell & Grance, 2009). Some of the benefits (Grossman, 2009) of Cloud Computing are scalability, simplicity of implementation and lower capital expenditure. While industry leaders and customers have wide-ranging expectations of benefits from Cloud Computing, privacy and security concerns remain a major impediment to widespread adoption (Westervelt, 2009).

In this exploratory study, we surveyed 61 Cloud Computing vendors about their information assurance (IA) practices. The data for this study has been collected from statements about privacy policy, acceptable use policy, terms of use and service level agreements available on the websites of the Cloud vendors. In the case of any such statements missing from the websites, similar information was sought via Internet searches of whitepapers, press releases as well as news articles of Cloud Computing in IT magazines. We also presented *emphasis index* and *assurance information availability index* for each information assurance practice.(Chakraborty et al., 2009) The data is analyzed using both descriptive and inferential statistics (ANOVA) against a few important attributes of Cloud vendors. We focus on three categories of Cloud Computing services, the size of the Cloud vendor as well as its familiarity in the IT world. The three well-known broad categories are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). Online traffic for the vendor sites (site visits) have been used as a proxy for vendor reputation. The paper demonstrates some areas where significant differences are found and others where there are no significant differences across different dimensions of information assurance practices. In a previous similar study (Chakraborty et al., 2010) with a much smaller sample of Cloud Computing vendors we came up with trends of IA practices based on different vendor characteristics. In this paper, through ANOVA and with a much larger sample set a more thorough analysis has been performed to investigate possible differences among vendors in terms of IA practices and their characteristics.

BACKGROUND:

Several important Cloud Computing initiatives have been launched in recent years to address IA practices. Cloud Security Alliance (CSA) is an organization that was launched in April 2009 to raise awareness about Cloud Computing security issues. CSA has since released a report on security guidance for critical areas of focus in Cloud Computing (Westervelt, 2009). The Open Cloud Manifesto is another initiative committed to making the Cloud as open as possible and to guide the Cloud Computing community to adhere to a core set of IA principles(2009; manifesto, 2009). The National Institute of Standards and Technology (NIST) created a Cloud Computing security group to determine the best way to provide security for agencies that want to adopt Cloud Computing (Balding, 2009). The World Privacy Forum Report on Cloud Computing security frames and analyzes the issues of privacy and confidentiality in this environment(Gellman, 2009).

Types of Cloud Services

While there is a lack of standardization, several IT experts (Vaquero et al., 2009; Networks, 2009) have classified Cloud Computing vendors into 3 broad categories based on the fundamental nature of the Cloud-based solution they provide. These are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) (Weinhardt et al., 2009). In the following paragraphs as well as in Figure 1 we explore each of the proposed groups of Cloud services (Lin et al., 2009) and solutions.

Infrastructure-as-a-Service (IaaS):

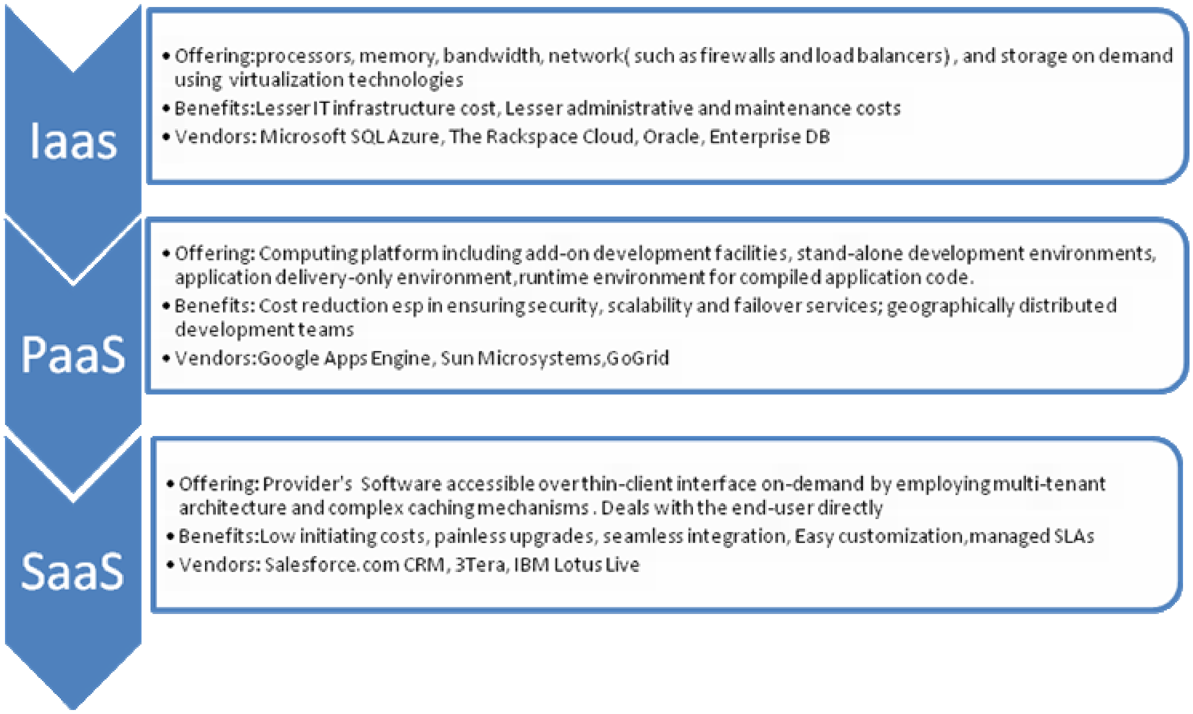
When a vendor rents infrastructure components such as processors, memory, bandwidth, network, and storage on demand, it is said to be delivering an IaaS service. The rented computational resources are the core components of IT services infrastructure in an enterprise. This service is facilitated primarily by advances in virtualization technologies and can be used as a platform for building applications or to enable entire applications. A big advantage of the IaaS approach is the possible elimination of capital expenditures on IT infrastructure from the client's side. IaaS services could be further categorized as Hardware-as-a-Service (Amazon AWS), Database-as-a-Service (Oracle, EnterpriseDB), and Storage-as-a-Service (Amazon S3).

Platform-as-a-Service (PaaS):

PaaS delivers a Computing platform and solution stack as a service. PaaS facilitates the deployment of customer-created applications to the Cloud using provider-supported tools (e.g., java, python, .Net). While the consumer does not control the underlying Cloud infrastructure, it has control over the deployed applications and hosting environment configurations(Mell & Grance, 2009).

Software-as-a-Service (SaaS):

In this type of Cloud, a provider's specialized software runs on a hardware Cloud infrastructure and is accessible to the customer through a thin client interface such as a web browser. Customers may configure application settings according to their specific needs. Some examples of online applications that SaaS providers employ are Clarizen project management tool, Customer Relationship Management (CRM) and human-resource applications by Salesforce, and project lifecycle management service by Absolute Performance. Certain SaaS solutions are essentially desktop applications that access and manipulate data hosted in the Cloud. A complex caching mechanism distributes the computation across the customer's computer, the Internet and remote data centers so that customers are unaware of the location of computation(Swaminathan, 2008). The figure below outlines the differences between IaaS, PaaS and SaaS in regards to offering, benefits and Vendors



- Offering:processors, memory, bandwidth, network(such as firewalls and load balancers) , and storage on demand using virtualization technologies
- Benefits:Lesser IT infrastructure cost, Lesser administrative and maintenance costs
- Vendors: Microsoft SQL Azure, The Rackspace Cloud, Oracle, Enterprise DB

- Offering: Computing platform including add-on development facilities, stand-alone development environments, application delivery-only environment,runtime environment for compiled application code.
- Benefits: Cost reduction esp in ensuring security, scalability and failover services; geographically distributed development teams
- Vendors:Google Apps Engine, Sun Microsystems,GoGrid

- Offering: Provider's Software accessible over thin-client interface on-demand by employing multi-tenant architecture and complex caching mechanisms . Deals with the end-user directly
- Benefits:Low initiating costs, painless upgrades, seamless integration, Easy customization,managed SLAs
- Vendors: Salesforce.com CRM, 3Tera, IBM Lotus Live

Figure 1: Different Types of Cloud Services

INFORMATION ASSURANCE

There are certain minimum information practices that Cloud Computing vendors should follow to ensure the confidentiality, integrity, and availability of customer's data. The issues of end-to-end security and privacy are of greater complexity in a Cloud Computing world than within a single data center. Some key issues of Cloud security include trust, multi-tenancy, encryption and compliance(Mell & Grance, 2009). The potential risks of using Cloud services include loss of direct control of resources and increased liability risk due to security breaches and data leaks due to shared external resources. Additionally, reliability loss is a distinct risk since service providers may go out of business, causing business continuity and data recovery issues (Mather et al., 2009). Despite the potential information assurance risks, Cloud Computing also offers some distinct benefits. For example, shift of public data to an external Cloud reduces the vulnerability of internal sensitive data to exposure. Centralized resources provide a level playing field for businesses of all sizes. Redundant data storage in offsite locations can also improve disaster recovery capabilities(Buyya et al., 2009). We shall explore business continuity and integrity at a later stage of this research.

The objective of our investigation is to identify the establishment of security and privacy practices in different groups of Cloud Computing providers and the degree of emphasis given by each provider.

Information Assurance Dimensions:

Information Assurance requirements for Cloud providers may be grouped under three dimensions described below:

(1) Security

There is a high possibility of security threats in Cloud IT infrastructures as vendors store critical and confidential data in the Cloud. In some cases, there is a requirement from the clients for physically or virtually separating data and applications. While Cloud providers can invest in better security controls through scale economies, they can also develop standardized processes for regulatory compliance. While it is evident that Cloud providers endeavor to improve their offerings to meet clients' enterprise-grade security needs, it may not be enough in some key sectors. In sectors such as defense, aerospace and brokerage, security and compliance requirements which include the physical location of the data have made SaaS and Hardware public clouds unacceptable for a while(Swaminathan, 2008). In a recent survey, 64% of respondents in the US Federal Government say security is their topmost concern in the context of Cloud Computing (Chabrow, 2009). The trust levels towards Cloud Computing in these sectors have however been improving. The launch of the Federal Cloud services portal for government agencies called Apps.gov is indicative of this shift. Vendors like Google and Microsoft are close to obtaining accreditation for FISMA (Federal Information Security Management Act) compliance for its Cloud-Computing services in order to be acceptable for the public sector(Howard, 2009). Recently Microsoft recommended US Congress to pass the "Cloud Computing Advancement Act", that also calls for an update to the Electronic Communications Privacy Act and the Computer Fraud and Abuse Act(Hoover, 2010).

(2) Privacy

Cloud Computing clients have major concerns regarding privacy since they don't know if the third-party or the Cloud Computing vendor has a privacy policy similar to or better than their own. As a basic requirement, the Cloud Computing client should be given the flexibility to assign an access control list on how, when and whom to be given the access to its data on the Cloud. A number of Cloud Computing clients desire to have access to access-logs and audit trails of all users and provider's employees. Additionally, Cloud vendors may need provisions for external audits on their infrastructure and network (Wang, 2009), and open standards identity management systems (Lin et al., 2009).

ATTRIBUTES OF CLOUD VENDORS:

For this study, the first 80 Cloud Computing vendors were chosen from a list of 150 top Cloud Computing vendors that appeared in the Cloud Computing Journal on Oct 29, 2009. It must be noted that this list did not put forth the names of the vendors in any particular order.

Previous studies (Kim et al., 2004; Sivasailam et al., 2002) have demonstrated significant differences in information assurance practices by B2C websites that provide information as well as physical goods based on the industry type. Different industries have different types of clients whose aspirations about information assurance are different as well. Similarly clients accessing different kinds of Cloud Services might have different priorities for information assurance practices. Hence, although Cloud Computing is not restricted to B2C transactions (e.g. Salesforce provides SaaS solutions to businesses, not consumers), it is possible that different Cloud service types will bring out significant difference in IA practices just as we observed for B2C websites. Thus we chose Cloud service type as a key attribute of the vendor for our comparison purposes. Reputation

We thus compared the Cloud vendors with respect to (a) the type of Cloud service(s) they offered (IaaS, PaaS or SaaS) and (b) company size. The size of the company is measured by the number of employees that are currently on payroll and this information was available from the social networking website called Linkedin. We believe this measure is quite appropriate in the context of an online delivery model like Cloud Computing since a large vendor may dedicate disproportionately low number of employees to their Cloud services department and the policies could be dictated by people from outside that department. Furthermore, at this stage we have not been able to define the boundaries of a Cloud service department and some of the Cloud-related transactions like payment and customer service might involve people outside such a department. We have divided the company sizes into reasonable groups based on the classification commonly used in the US: (i) "Small" for number of employees below 100, (ii) "Medium" for number of employees between 100 and 500, and (iv) "Large" for number of employees larger than 500. It should be noted that in the case of some large vendors, this measure does not capture the size of the employee pool involved in Cloud-only services. In the absence of more accurate measure, we believe that size of the employee pool is a reasonable proxy for capturing Cloud vendor size. In order to explore the characteristics of the Cloud vendors, we present the following hypotheses:

H1: There is a significant difference in information assurance practices with respect to security and privacy between Cloud Computing vendors based on the type of Cloud services they offered.

H2: There is a significant difference in information assurance practices with respect to security and privacy between Cloud Computing vendors based on the category of company size they belonged to.

DATA COLLECTION

A questionnaire (available on request) with two dimensions – security and privacy was used to collect available information from policy statements, terms of use, and acceptable use policy as well as news articles for 61 Cloud Computing vendors and providers. To measure security, in addition to secured authentication and secure data transfer, we analyzed whether Cloud vendors comply with different information assurance regulations of the US like HIPAA and SOX. Privacy questions are geared to discern if vendors grant flexibility to users on data control. Some of the vendors for our data set are presented in

Table 1 with their corresponding Cloud Service Type and Vendor Size Group. Not all vendors have been included in this table for length constraints.

Cloud Vendor	Cloud Service Type	Vendor Size Group
Amazon Web Services	I,P,S	Large
Google	P,S	Large
IBM	I,P,S	Large
Salesforce.com	I,P	Large
Oracle	I,P,S	Large
Unisys	I,S	Large
NetSuite	P,S	Large
Parallels	I,S	Large
Rackspace	I,P	Large
Opsource	I	Medium
Appirio	S	Medium
Platform Computing	I	Medium
Meeza	S	Medium
eVapt	S	Medium
Layered Technologies	I	Medium
Workday	S	Medium
Wyse	S	Medium
Elastra	I,S	Small
LongJump	P,S	Small
rPath	S	Small
Engine Yard	P	Small
Enomaly	I,S	Small
ThinkGrid	I,S	Small
Parascale	P,S	Small
Skytap	I,P,S	Small
3Tera	I,P,S	Small

I=IaaS, P=PaaS, S=SaaS, Small: <100, Medium:100-500, Large: >500

Table 1: Cloud Computing Vendors, their Service Types and Size

MEASUREMENT

In order to measure the responses to the questions, we first coded them according to the following scheme: a positive response to any question was assigned a 1; a negative was assigned -1; the inability to answer any question either way was indicated with a 0. Since the questions in our survey instrument represented different aspects of each dimension of information assurance, the measurement would reflect the “extent” of emphasis a Cloud vendor or provider places on the corresponding dimension of information assurance. We adapted the Emphasis-Density Index measure of (Kim et al., 2004) which is equal to the positive (or negative or N/A) number of items in a survey for each dimension divided by the total number of items.

Three indices are measured: (a) Positive Emphasis Density Index, (b) Negative Emphasis Density Index and (c) Assurance Unavailability Index. Each of these indices is applied to each dimension of information assurance. A positive index of a certain dimension is measured by counting the number of 1's found in the responses to the questions on that dimension which is then divided by the total number of corresponding items, e.g. 8 for security. Similarly, for the negative index we counted the number of -1's while for the unavailability index we measured the number of 0's.

RESULTS AND ANALYSIS

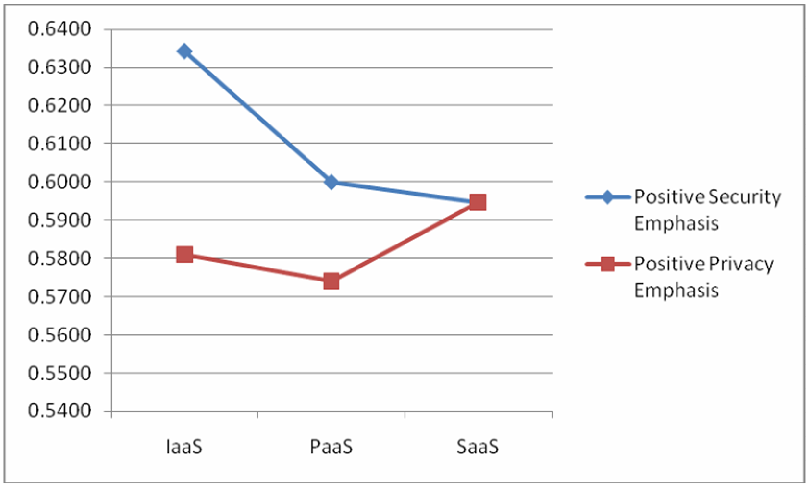


Figure 1: Positive Emphasis Indices for different types of Cloud Services

After conducting ANOVA for testing both hypotheses (H1 and H2), we saw a partial support for the effect of Vendor Size on the security practices. In particular, statistically significant difference was found in the Positive as well as Negative Security Indices between the different size categories (Large, Medium and Small) of Cloud vendors. In combination with the plots of the descriptive statistics we could surmise that medium sized vendors actually have worse practices in Information Assurance compared to the small sized ones. This is perhaps one of the biggest symptoms of why Cloud Computing has not been embraced completely. No other hypotheses were supported from our studies. From the remaining of the findings where especially no significant difference of IA practices was found between different Cloud service types, we should take the hint that most vendors are moving towards a homogeneous kind of assurance and many of them are indeed serving up multiple types of services. In future we should thus try to understand better the nuances among services and come up with a better taxonomy of Cloud solutions and services.

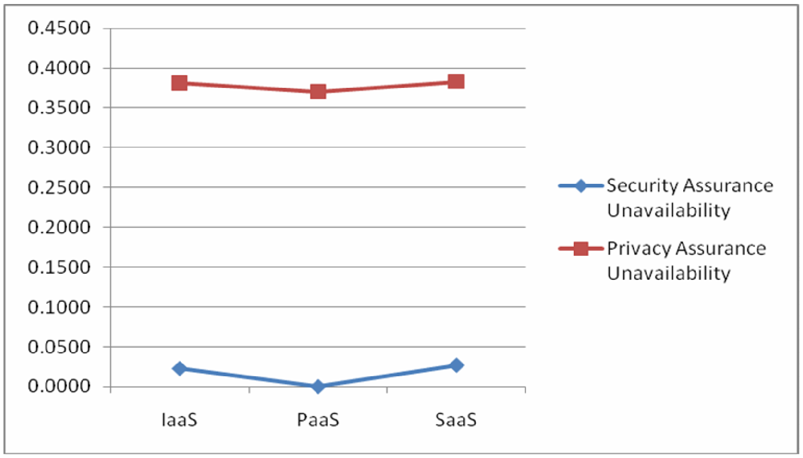


Figure 2: Assurance is less concrete when it comes to privacy practices of the Cloud vendors

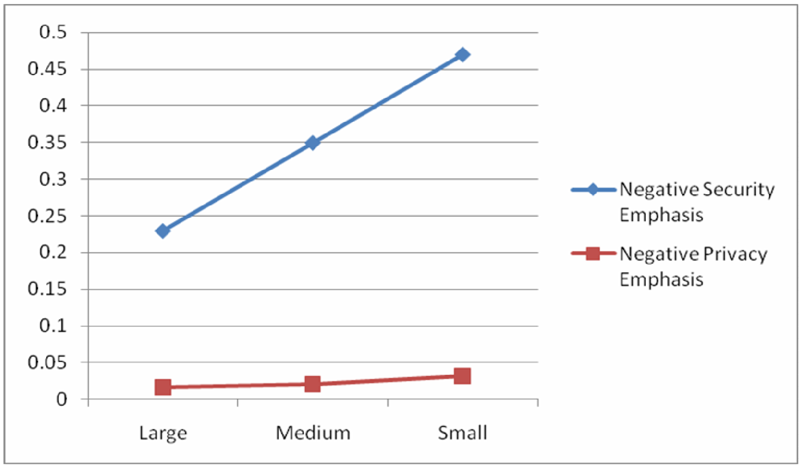


Figure 3: The smaller the company is, the less are their assurance practices

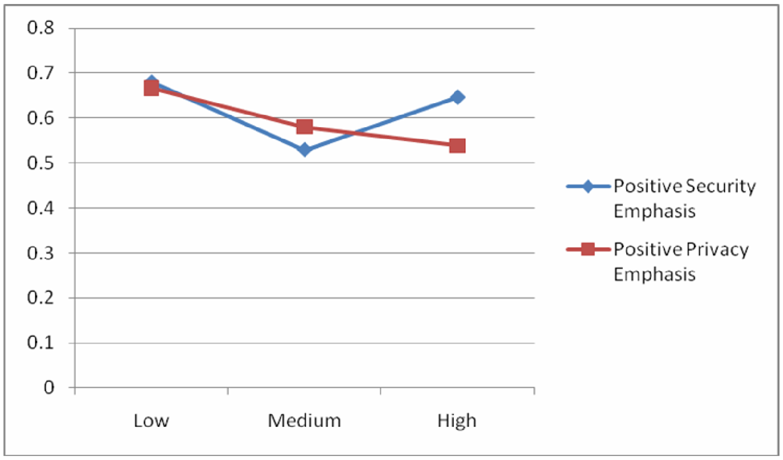


Figure 4: In this odd finding, the lesser known companies seem to put a higher emphasis on security

FUTURE WORK

In future, we would like to investigate if the degree of complexity of Cloud-based solutions and services has any effect on the importance that the Cloud vendor gives to information assurance. For example, it will be useful to see how the emphasis on IA changes when the Cloud solution changes from a combination of PaaS and IaaS to simply IaaS. Many vendors will offer both the combination as well as the single service and this implies that a more accurate assessment of IA practices based on Cloud Service Types requires treating them separately. In addition to the Cloud Service Type it will be also important to look at IA practices with respect to different categories of solutions within the same Cloud Service Type. For example, it is important to see if there is a significant difference in IA practices between a vendor that offers CRM as a SaaS solution (Salesforce) and a vendor that offers netbook operating systems as a SaaS solution (ChromeOS from Google). If a difference is observed, it will be imperative to compare that with the differences between vendors offering equivalent but non-Cloud-based solutions and products. We have also noticed that a lot of these vendors are increasingly providing consultancy for a comprehensive Cloud-based solution. For example, Accenture is now offering a “Cloud Computing Accelerator” which itself is not a Cloud-Computing service but rather overall package that starts with assessment about moving some IT infrastructure to the Cloud followed by some actual SaaS implementation. This brings us to an additional attribute of the Cloud vendor to investigate with respect to a vendor’s IA practices. We foresee that the more “complete” the solution is the better will be the emphasis index scores that we have used in this paper. Finally, we would like to improve upon the current metric, i.e. the emphasis index and empirically establish some weights to combine them together to form a single metric as opposed to presenting three separate findings (positive, negative and unavailability) for each vendor attribute as we have done in this paper.

CONCLUSION

In this brief study, we investigated various Cloud Computing vendors (mostly from the US) by trying to get some key Information Assurance related questions answered from whatever information the vendors post on their websites. In addition to the directions identified in the previous section, in future the method of data collection employed in this study needs to be complemented by a thorough qualitative study of magazine articles, SEC filings of Cloud clients as well as interviews with executives from both vendors and clients of Cloud Computing. This comprehensive data collection aided with extant MIS theories will be used to synthesize theories about antecedents of IA practices in Cloud Computing. Nevertheless, studies like these have not been conducted enough given the pace at which Cloud Computing is picking up attention of the CIOs around the country. We thus hope that this is a promising beginning of a journey of looking into more sophisticated performance and strategy issues that interconnect with information assurance in the Cloud Computing.

REFERENCES

2009. http://www.opencloudmanifesto.org/. In O. Cloud (Ed.).

Balding, C. 2009. US Government Creates Cloud Computing Security Group *cloudsecurity.org*.

Behnia, K. 2009. Cloud Computing: Time to Rethink IT Service Delivery and Bring the Clouds Down to Earth.

Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., & Brandic, I. 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems*, 25(6): 599-616.

Chabrow, E. 2009. Rules Make Adoption of Cloud Computing Challenge for Agencies.

Chakraborty, R., Ramireddy, S., Raghu, T.S., & Rao, H.R. 2009. *A Comparative Study of Cloud Computing Vendors & their Information Assurance Practices*. Paper presented at the Proceedings of 3rd International Conference on the Virtual Computing Initiative.

Chakraborty, R., Ramireddy, S., Raghu, T.S., & Rao, H.R. 2010. An Exploratory Study of Information Assurance Practices of Cloud Computing Vendors. *IEEE IT Professional*(Forthcoming).

Gellman, R. 2009. Privacy in the Clouds:Risks to Privacy and Confidentiality from Cloud Computing. *World Privacy Forum*.

Grossman, R.L. 2009. The Case for Cloud Computing. *IT Professional*, 11(2): 23-27.

Hoover, J.N. 2010. Microsoft seeks new legal framework for cloud. *Information Week*.

Howard, A.B. 2009. FISMA compliance for federal cloud computing on the horizon in 2010. *SearchComplaince.com*.

Kim, D.J., Sivasailam, N., & Rao, H.R. 2004. Information Assurance in B2C Websites. *Electronic Markets*, 14(4).

Lin, G., Fu, D., Zhu, J., & Dasmalchi, G. 2009. Cloud Computing: IT as a Service. *IEEE IT Professional*, 11(2): 10-13.

Vaquero, L.M., Rodero-Merino, L., Ceceres, J., & Lindner, M. 2009. A Break in the Clouds: Towards a Cloud Definition. *ACM SIGCOMM Computer Communication Review*, 39(1): 50-55.

Mather, T., Kumaraswamy, S., & Latif, S. 2009. *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance*: O'Reilly Media, Inc.

Mell, P., & Grance, T. 2009. Effectively and Securely Using the Cloud Computing Paradigm.

Networks, F. 2009. *F5 Study Shows Cloud Computing Gaining Critical Mass Among Large Enterprises*.

Swaminathan, K.S. 2008. Computing in the clouds. *Outlook Journal*.

Sivasailam, N., Kim, D.J., & Rao, H.R. 2002. *What Companies Are(n't) Doing about Web Site Assurance. IEEE IT Professional*, 4(3): 33-40.

Wang, C. 2009. A close look At Cloud Computing Security Issues. *Forrester*.

Weinhardt, C., Anandasivam, A., Blau, B., & Stößer, J. 2009. Business Models in the Service World. *IEEE IT Professional*, 11(2): 28-33.

Westervelt, R. 2009. Cloud computing security group releases report outlining trouble areas, *SearchSecurity.com*.