JOURNAL OF SOFTWARE ENGINEERING & INTELLIGENT SYSTEMS
ISSN 2518-8739
31st August 2016, Volume 1, Issue 1, JSEIS, CAOMEI Copyright © 2016
www.jseis.org

Risks with cloud computing data residency and virtualization

Zubair Ahmad[1], Akram M. Zeki[2], Akeem Olowo[3]
Faculty of Information & Communication Technology, International Islamic University Malaysia, Malaysia.
Email: [1]zubair8vaani@gmail.com, [2]akramzeki@iium.edu.my, [3]akeemolowo@gmail.com

ABSTRACT

Cloud computing helps organizations to shift focus from technology challenges to improving business performance. Cloud computing is widely adopted across many industry sectors. However, with adoption comes a security concern. Many approaches have been adopted or recommended to ensure privacy and prevention of data loss. Apart from security concerns of cloud computing, there are more other challenges like issues surrounding data residency in cloud computing, especially with regards to where the data is stored and what laws apply in global computing environment, legal issues involved in different countries may take weeks or months or even years to get sorted out? Further concerns, include knowing when data is breached, or if data remain in the cloud even after termination of service or if data is seized, making the business to face threat of availability by not being able to continue operations. In this paper we discuss broadly the risks associated with cloud computing data residency and virtualization and propose assessment solutions to secure the data.

Keywords: data residency; virtualization; data security; encryption; tokenization techniques; virtualization techniques;

1. INTRODUCTION

Cloud computing is experiencing significant growth with rapid adoption in various regions around the world. By adopting cloud technology, lots of companies have managed to reduce total cost of ownership, increased flexibility in IT implementation, become more competitive amongst emerging players and meet time to market objectives. Cloud technology provides business benefits. With large scale data centers connected with the high speed internet networks, the benefits of cloud computing are huge, but however there are few challenges associated with cloud computing like data residency, data security and risk imposed by new technologies. Such concerns include virtualization introduced in cloud world. One of the questions that bother data and security experts is the consideration regarding where in the world the data is located? Another intriguing concern is the issue of application of legal consequence for any infringement on data. For instance, when one travels, any legal issue that arise is resolved with the law where you are and not by the country you come from if you run into problems you are subjected to that country laws to sort things out you need to deal with local authorities. Traditional boundaries of state and country have been blurred by the internet, e-commerce and especially cloud computing. Cloud computing transactions between companies often involve companies from one country, procuring cloud computing service in another country and probably customers in a third country. Offshore cloud services modes of resolving disagreements and imposing judgements requires different approaches than the conventional means. For instance, the bond may be signed in Australia, the supplier are the local subsidiary of US company and service located in Malaysia. This raises the prospect of very expensive international derogation. Hence, it is imperative before agreeing on cloud computing services' terms and conditions to have a clear understanding of the services, especially knowing where the data resides. Furthermore, if data is seized or compromised, the business faces the threat of not being able to operate. Consequent upon which legal issues may ensue. However, when it involves different countries, legal issues resolutions may take weeks or months or even years to get sorted out.

2. DATA RESIDENCY AND DATA SECURITY ISSUES

The main concern with data residency is who manages and has access to the data where the data is stored and what laws apply? How to know when data is breached? Will data remain in the cloud even after termination of service? From the aforementioned, obviously there are many risks involved in data residency. One of the risks involved in cloud computing is that the data that resides in cloud may be available to more people and services that are managed by third parties thereby there is the possibility that someone from those third parties might have access to the data.

Dealing with cloud computing and considering that data is extremely important or sensitive, companies want to consider making sure they put limits on people to see or maybe decide not to put their data into the cloud or maybe encrypt their data for protection in the cloud. These are decisions taken to mitigate risks. and allow that particular risk in the environment. Another challenge from security perspective is that security of clouds is

managed by another company elsewhere and the control mechanisms are in the hands of a third party. For instances, in the case of companies such as YahooMail or GoogleMail, a typical user doesn't manage or handle the security by themselves. They simply trust Google or Yahoo with security and believe for certain that their email accounts are secure and nobody else gets access to users' accounts except when users themselves compromise their security. Security breaches often occurs on the side of the providers themselves and when such occurs, that could have very damaging or devastating effect on the users. Using this scenario, if the situation is likened to companies with data residency in the cloud, it becomes a bit of a challenge because data is being entrusted in the hands of third parties. Hence, if company's data and information reside in the cloud, managed by a third party, then some of these concerns are legitimate and all the pros and cons should be certainly considered. Another issue of important concern is considering that cloud computing servers are somewhere else and since the service and facilities belong to a third party hence there is limited control on whatever happens on the server. For instance, if the server goes down or in the unlikely event of power outage or due to system or storage failures or technical lockout of account, there may not be direct access to fix these particular issues. Employing Services in the cloud doesn't mean that these services are always available. Likewise, these systems are managed by humans and sometimes what happens there on cloud creates downtime. These are the issues to be kept in mind since these poses some availability risk for the organization not having access to the systems [1] [3] [4].

- Another concern is data residency laws. Do local data centers solve the data residency issue? Europe in particular have strict laws around data residency and in recent times, many US companies have started opening data centers in European countries. However, the pertinent question is does this solve the problem and what are the likely issues that may emerge from this approach? Some of the issues are involved here as cloud providers' open data centers across multiple countries and multiple regions having different laws include the fact that a particular company's data may not be guaranteed to stay in one data center This actually exacerbated in that most cloud providers back up data or recommended that you to make backup of your data across multiple regions. Providers typically have some commands and controls to access data even if it is stored in European data center [2] [5]. Characteristic strategies of cloud providers in controlling data across borders include Regional data centers are rarely autonomous.
- Data centers are mainly used for redundant backups in the US.
- Centralized "Command and control" are usually used to communicate data which can be in any country.

According to Gartner the core part of security deliverables are the compliances. In a given company, cloud clients, IT divisions, and CIOs/CISOs all offer obligation regarding compliances. Gartner in its research interviewed more than 200 IT pioneers in organization with the Cloud Security Alliance and found that 21% of organizations have a cloud administration board of trustees in charge of building up and authorizing arrangements that ensure corporate information in the cloud. IT security is the most well-known groups found on these panels, trailed by IT management, cloud computing law, and compliance/risk groups. Considering the budgetary penalties of a compliance infringement, and the consequent rupture revelation, flood of client claims, more organizations likely need a formal procedure for compliance that incorporates an expansive arrangement of stakeholders [10].

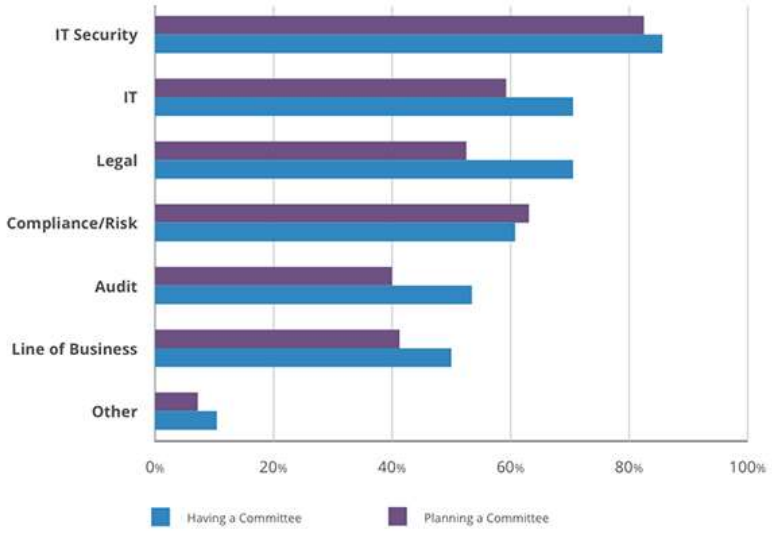

Figure. 1 IT Divisions Comparison

3. RISKS ASSOCIATED WITH CLOUD VIRTUALIZATION:

Another technology that has become widespread is virtualization. This is the idea of having one big single server and built virtual systems. Unlike in the past, a company's capacity may require 20 different servers however now one big server can virtually be used to create these 20 servers all residing inside of this one server. The nice this arrangement is that it affords the opportunity to have lot of control over what can take place with that system. Besides, there's the ability to add more memory, or increase disk space thereby not limited by physical constraints anymore. This provides lots of business values associated with virtualization.

But from a security perspective, there are lot of emerging set of threats associated with taking advantage of virtualization layer. Virtualization layer is the layer which sits on top of the virtual machines. Hackers typically target this layer. Hackers or someone with wrong intentions know if they compromise that virtualization layer then there is a potential chance of gaining access to every single virtual system that might be on that physical server. Imagine having hundreds of virtual machines on a single server and risk the company's cloud infrastructure presents to not only servers but customers data may also be at risk [6] [7] [13]. These are the challenges with virtualization that cannot be simply ignored.

Furthermore, there is less control over virtual machines to virtual machines communications and even not much support for "virtual firewalls". There are also challenges when one start looking at multiple systems cramped into one physical device. Obviously in a data center if it was a physical server there is a lot of control over who access the servers physically. There is also the ability to even separate these servers off in a completely different physical area of the data centers and in some cases in separate data centers. This arrangement presents some advantages for being able to separate server into different physical environment from data perspective and physically. However, in virtual environment all servers are stuck together into one system. Separation of system security becomes a little bit harder to manage. It is possible to manage separation in virtualization with some software tools but it is necessary to make sure they are implemented properly such that different systems are moved into different VLANS, that is, since physically different virtual systems cannot access each other though implementation is not that easy compared to physical separation.

Again, from a business management perspective it is also necessary to be clear about separation of duties when everything is on one single big server. For instance, if all of company's databases are in separate different virtual machines on a single server, separation of duties may become more difficult to administer. This becomes more complicated to managing one single server that contains these different virtual machines. Consequently, there is the need to have proper policy n managing a virtual server, more especially if there are multiple people that are managing that virtual server. Other possible scenarios include having the administration of that server split off into other pieces or maybe there is an overlay on top of every single virtual server for management and security [8] [11].

4. DATA ENCRYPTION AND TOKENIZATION SOLUTION TO PROTECT DATA

Data encryption and tokenization are the key solutions to overcome data residency concerns. Data encryption is the mathematical process of converting data in *plain text* data into *cypher text* which cannot be read by any entity other than entity(ies) that retains the valid encryption key. In tokenization, the actual data resides locally in a token database where randomly generated tokens are associated with the data and are sent to the cloud. The data can only be read by the custodian of the token database [14-16].

The key differences between *tokenization* and *encryption*. *Tokenization* protects only against the external threats since anyone with access to the token database could access clear text data whereas *encryption* protects against internal and external threats since there is segregation of duty between where the keys are managed and where the encrypted data is stored. Tokenization requires higher capacity servers and databases and more operational oversight to manage the critical token database that grows with increasing data volumes. Encryption requires lighter weight stateless servers with *No data storage*. This helps organizations to shift focus from technology challenges to improving business performance [9] [10] [12].

5. CONCLUSION

In this paper we discussed the major issues related to data residency concept of information in the cloud and how risky it can be for small companies to pin down and understand where their data or customers' data is actually located or stored. Generally, one of the very first things done by procuring companies is to look at contract terms and need to determine what jurisdiction their data is in. A number of global cloud providers have standard contract based in most of cases on the US law which requires cloud providers to disclose where data will be stored or whether multiple jurisdictions are involved, in which case small or medium companies must take legal advice in

every jurisdiction were their data could end up. Associated with data residency is significant amount of security risks especially due to virtualization. These threats are posed by someone taking advantage of the virtualization layer since hackers are aware that access to virtualization layer presents a potential to gaining access to every single virtual system that might be on that physical server. Finally, an attempt was made to discuss data protection techniques for clouds mainly encryption and tokenization which is meant to secure data and ensure privacy whether the cloud sits locally or globally.

ACKNOWLEDGMENT

The authors would like to thank the Research Management Center and the faculty of Information and Communication Technology at the International Islamic University Malaysia for their support.

REFERENCES

1. Badger, L., Grance, T., Patt-Corner, R., & Voas, J. (2011). Draft Cloud Computing Synopsis and Recommendations. National Institute of Standards and Technology (NIST) Special Publication 800-146. US Department of Commerce. May 2011.
2. Md.T. Khorshed, A.B.M.S. Ali, S.A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing", Future Generation Computer Systems, 28(6), June 2012.
3. D. Zissis, D. Lekkas, "Addressing cloud computing security issues", Future Generation Computer Systems, 28 (3), March 2012, Pages 583-592
4. Y. Mei, L. Liu, X. Pu, S. Sivathanu, "Performance measurements and analysis of network I/O applications in virtualized cloud", CLOUD 2010
5. P. Mell, T. Grance, The NIST Definition of Cloud Computing, NIST Special Publication 800-145, 2011.
6. M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, On Technical Security Issues in Cloud Computing. IEEE, 2009.
7. Cloud security alliance: Security guidance for critical areas of focus in cloud computing v2.1," Dec 2009.
8. Sándor Ács, , Miklós Kozlovszky; Advanced Vulnerability Assessment Tool for Distributed Systems; HP-SEE User Forum 2012, BoA. pp. 46,Belgrade, Serbia, October 17-19, 2012
9. M. Kozlovszky, M. Törőcsik, T. Schubert, V. Póserné; IaaS type Cloud infrastructure assessment and monitoring , MIPRO 2013 may, Opatija, Croatia
10. SN Chiueh, "A survey on virtualization technologies", RPE Report, 2005
11. Bertion, E., Paci, F., & Ferrini, R. (2009). Privacy-Preserving Digital Identity Management for Cloud Computing. IEEE Computer Society Data Engineering Bulletin, pp. 1-4, March 2009.
12. Center for the Protection of Natural Infrastructure (CPNI)'s Information Security Briefing on Cloud Computing, 01/2010, March 2010.
13. "Virtualization and Cloud Computing Threat Report." Trend Micro. August 2011.
14. Brian O. and others, Cloud Computing, authors:,2012-11- 06, page 6, publish Swiss.
15. Sehgal NK, et al.: Information Security and Cloud Computing, Iete Technical Review, Vol 28, Issue 4, Jul-Aug 2011
16. Rajiv R.Bhandari, Mishra N., Encrypted IT Auditing and Log Management on Cloud Computing, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 1, pp. (302), September 2011

AUTHOR PROFILE

Zubair Ahmad is a master student in Information Technology at faculty of Information and Communication Technology, International Islamic University Malaysia, MALAYSIA. His research interests are in networking, communication and cloud computing.

Akram M. Zeki is an Associate Professor at faculty of Information and Communication Technology, International Islamic University Malaysia, MALAYSIA. His research interests are in networking, communication and cloud computing.

Akeem Olowolayemo is a post doctorate researcher at faculty of Information and Communication Technology, International Islamic University Malaysia, MALAYSIA. His research interests are in networking, communication and cloud computing.