Building Trust and Confidentiality in Cloud computing Distributed Data Storage

Uegebe Ikechukwu Valentine and Omenka Ugochukwu Enyinna
Department of Computer Science, University of Port Harcourt, Rivers State
Email:donnumerator2002@yahoo.com Tel: 08032634264; Email:uomenka@gmail.com Tel: 08037430135

Abstract

Cloud computing delivers massively scalable computing resources as a service with internet technologies. Resources are shared among a vast number of consumers allowing for a lower cost of IT ownership. Enterprises can store or rent data storage as a service in a "pay-per-use" manner. As with new technology, this new way of doing business brings with it new challenges, especially when considering the security and privacy of the information stored and processed within the cloud. In this paper, we looked at data security, described the current state of data security in the cloud and the possible threats obtainable in cloud computing. We described how the combination of existing research thrusts has the potential to solve many threats concerning confidentiality and adoption of cloud. We proposed that with continued research and adoption of trusted computing and computation-supporting encryption, maintaining integrity of data in the cloud will be a success.

Keyword: Cloud computing, Security, privacy, internet services, virtual machines**.**

1.0 Introduction

Information technology has become pervasive in organizations and an inevitable key success factor in business. Organizations can create, communicate and collaborate faster, more efficiently and reliably than ever before. In the late 1960's, the computer scientist John McCarthy once brought the concept of utility computing in which he postulated that life cycle of technology will not only stick as tangible products, but will indeed become products. As a matter of fact, he took the conceptual leap to predict that computer resources will be provided like nowadays water and electricity as a utility,i.e. as a service [2]. However in the last couple of years internet services offered online took on an even new dimension. Software is now capable of being offered online including big fast machines in someone else's data centre running an application that is accessed using a familiar web browser, although someone else owns the application. Cloud computing delivers massively scalable computing resources as a service with Internet technologies, resources are shared among a vast number of consumers allowing for a lower cost of IT ownership [7]. Cloud computing provides on-demand

computing resources dynamically, which allows companies to fundamentally change their information technology strategy.

As with any new technology, this new way of doing business brings with it new challenges, especially when considering the security and privacy of the information stored and processed within the cloud. This article examines these challenges and proposes unique solutions in building trust in a distributed Data storage to solve the core security problems of cloud computing.

Utility cloud computing allows users to rent *Virtual Machine (VMs)* from a service provider, placing an organization's sensitive data in the control of a third party [5][10]. We propose a management and security approach for utility cloud computing called the *Private Virtual Infrastructure (PVI)* that shares the responsibility of security of data in cloud between the service provider and client, decreasing the risk exposure to both. To address the core security challenge of distributed data in cloud computing where an information owner creates and runs a virtual environment on a platform owned by a separate service provider from the inside, we introduce a new approach for rooting trust in a cloud computing environment called the *information centric approach.*

2.0 Data Security.

Security has been a major problem in the real world, not only in the IT world. So taking information and making it secure, so that only yourself or a select few can see it, is obviously not a new concept. However, it is one that we have struggled with in both the real world and the digital world. In the real world, even information under lock and key, is subject to theft and is certainly open to accidental or malicious misuse. In the digital world, this analogy of lock-and-key protection of information has persisted, most often in the form of container-based encryption. But even our digital attempt at protecting information has proved less than robust, because of the limitations inherent in protecting a container rather than in the content of that container [6][9]. This limitation has become more evident as we move into the era of cloud computing: Information in a cloud environment has much more dynamism and fluidity than information that is static on a desktop or in a network folder, so we now need to start to think of a new way to protect information.

Before we embark on how to move our data protection methodologies into the era of the cloud, perhaps we should stop, think, and consider the true applicability of information security and its value and scope. Perhaps we should be viewing the application of data security as less of a walled and impassable fortress and more of a sliding series of options that are more appropriately termed "risk mitigation" [9]. Susan Morrow in an article on "data security in the cloud" said that she want s people to start to view data security as a lexicon of choices, as opposed to an on and off technology. In a typical organization, the need for data security has a very wide scope, varying from information that is set as public domain, through to information that needs some protection (perhaps access control), through to data that are highly sensitive, which, if leaked, could cause catastrophic damage which nevertheless need to be accessed and used by selected users. Computer technology is a form of the toolkit that we have developed since human prehistory to help us improve our lifestyle. If we can view data security as more of a risk mitigation exercise and build systems that will work with humans (i.e., human-centric) [6], then perhaps the approach we proffer for securing data in the cloud will be successful.

2.1 The current state of data security in the cloud.

At the time of writing this article, data storage in cloud computing is at a tipping point: It has many arguing for its use because of the improved interoperability and cost savings it offers. On the other side of the argument are those who are saying that data storage in the cloud cannot be used in any type of pervasive manner until we resolve the security issues inherent when we allow a third party to control our information. These security issues began by focusing on the securing of access to the datacentre's that cloud-based information resides in. However, it is quickly becoming apparent in the industry that this does not cover the vast majority of instances of data that are outside of the confines of the data centre, bringing us full circle to the problems of having a container-based view of securing data [9]. But we are not in any way inferring that data-centre security is not used or has been replaced.

Going back to our previous statement that security is better described as "risk mitigation," we can then begin to look at securing data as a continuum of choice in terms of levels of accessibility and content restrictions: This continuum allows us to choose to apply the right level of protection, ensuring that the flexibility bestowed by cloud computing onto the whole area of data communication is retained.

Susan Morrow, in one of her articles said that IT industry is beginning to wake up to the idea of content-centric or information-centric protection, being an inherent part of a data object. This new view of data security has not developed out of cloud computing, but instead is a development out of the idea of the "*de-perimerization*" of the enterprise [6]. She further stated that this idea was put forward by a group of Chief Information Officers (CIOs) who formed an organization called the **Jericho Forum**. The Jericho Forum was founded in 2004 because of the increasing need for data exchange between companies and external parties—for example: employees using remote computers; partner companies; customers; and so on. The old way of securing information behind an organization's perimeter wall prevented this type of data exchange in a secure manner. However, the ideas forwarded by Morrow about the Jericho Forum are also applicable to cloud computing. The idea of creating protection within the data object itself, allows the security to move with the data, as opposed to retaining the data within a secured and static wall (firewall). This simple but revolutionary change in mind-set of how to secure data is the ground stone of securing information within a cloud and will be the basis of this discussion on building trust in a distribute data storage in the cloud.

2.1.2 Identified Treats in Cloud Data centre. Hypervisor and Rootkit Malware

A new class of attacks has evolved around building malicious hypervisors and operating system rootkits that subvert the built in security measures of many operating systems. These malwares utilize a hypervisor or rootkit that allows them operate at a privilege level above that of the guest operating system (OS) or maintain root access to the system. The malware at the higher privilege level can then intercept system calls from a victim OS and modify the calls in a manner that thwarts the security mechanisms of the victim VM [4].

The malware can gain access to protected memory, intercept passwords or cryptographic keys, and perform a multitude of other malicious acts that the guest OS has no chance of defending against as it would be able to do on a physical machine. An example of this type

of attack is SubVirt created by a University of Michigan research team, which is essentially a *Virtual-Machine Based Rootkit (VMBR)*. SubVirt has been used to implement a phishing web server, a keystroke logger, a service that scans the target file systems system looking for sensitive files, and a defensive countermeasure that defeats a virtual-machine detector [10][4]. The Blue Pill attack is another example of this type of attack. The Blue Pill is an attack to virtualize a Windows operating system by installing a malicious hypervisor underneath the kernel that is theoretically undetectable even though the algorithm and code are publicly available. It avoids detection by trapping all attempts by the victim OS to determine it is in a virtualized environment and reporting fake information back to OS to make it believe it is operating normally [4]. The Blue Pill attack can be performed on already virtualized machine, thus nesting itself between the real hypervisor and the victim machine. By verifying the validity of the hypervisor and host OS, we can determine if any malware was present in hypervisor and OS at boot time; however, an infection after boot time may not be detected. For this reason, we use encryption of data to reduce the risk of data exposure.

Data Loss and Leakage

Enterprises are lot more concerned about data loss and leakage. The threat of data compromise is much greater in the cloud. There are many ways data may be compromised in the cloud including deletion or alteration of records without a backup, loss of or changing an encryption key that results in the effective destruction of any data stored with the key, and unauthorized access by insiders or other cloud users. Again, encryption of sensitive data reduces the exposure of data loss and leakage.

Another area of vulnerability of the VM is while the VM is at rest (*i.e.* inactive) [4]. A VM that uses a virtual file system – as opposed to a physical one – is susceptible to data modification while the system is at rest. It is possible for an attacker to modify the configuration of the VM by manipulating the virtual file system and alter the behaviour, properties, and data stored on the VM. If an attacker gains access to a virtual file system, the data are vulnerable to theft as the attacker has full access to all data contained in the file system. Additionally, encryption of data in the VM image with keys locked to specific platforms reduces the exposure of ***data at rest attacks*** and data loss.

Malicious Insiders

A malicious insider is anyone in the service provider's organization that possess authorized access or privilege to the cloud information systems that is moved to compromise information confidentiality, integrity, and availability [4][3][5]. The insider threat is compounded when combined with lack of transparency into service provider processes and procedures. There is often little or no visibility into the hiring practices for cloud provider employees. For example, a provider may not reveal how it monitors employees or grants access to physical and virtual assets. Depending on the access granted, an insider could collect confidential data or even gain control of the cloud services with little or no risk of detection.

There are several attacks against the VMs that can be performed by malicious actors inside the *Cloud Virtual Fabric (CVF*). A malicious administrator can secretly attack a VM in the cloud in a way that no one can notice using her higher privileged access to inspect memory, monitor VM communications, and perform suspend and reboot attacks [4].

This attack is very difficult to defend against as the insider needs to have these privileges to administer and maintain the host systems and it is difficult to determine legitimate access versus malicious access. Therefore, the confidentiality and the integrity of the data would be violated when an adversary controls a node or the node administrator becomes malicious. Encryption of sensitive data reduces the exposure.

Network-Based Attacks

Virtual machines are vulnerable to network-based attacks, especially during attestation and live migration. These network attacks that can be performed include eavesdropping, man-in-the-middle, data modification, spoofing, etc. [4]. It is imperative that the network communication be thoroughly understood and examined to understand all the possible attacks against it. Most approach does not provide any direct protection from network attack; but encryption protocols do use cryptographic protocols which limit the exposure to network attacks.

3.0 Our Approach to building Trust and Confidentiality

There are at least two concerns when using the cloud; one concern is that

1. Users do not want to reveal their data to the cloud service provider. For example the data could be sensitive information like medical records. Another concern is
2. Users are unsure about the integrity of the data they receive from the cloud, therefore within the cloud more than conventional security mechanisms will be required for data security.

Yu Chen et al in his article ***secure distributed data storage in the cloud*** presented technologies for data security in the cloud computing from four different perspectives; [6]

i) Database outsourcing and query integrity assurance
ii) Data integrity in untrustworthy storage
iii) Web- application based security
iv) Multimedia data security storage.

All the technologies mentioned by Yu Chen are all effective, but will be more effective when being viewed from an information-centric security approach. Let us quickly look at what information centric security means.

Information-Centric Security

For us to maintain trust and confidentiality and extend control of data in the cloud, we propose *shifting from protecting data from the outside to protecting data within*. This is referred to as **information-centric**. This implies protecting the data content itself. Data needs to be encrypted and packaged with usage policy [10][1]. Information centric security is a natural extension of the trend towards finer, stronger and more usable data protection [3].

In our vision, we propose the use of trusted computing which ensures integrity of cloud infrastructure and in addition to the use of cryptographic protocols supporting computation on cipher text. Specifically, dual encryption approach is recommended for data object to a distributed data in the cloud [1][10]. This will enable cross examination of the outsourced data, which consists of

(a) the original data stored under a certain encryption scheme and

(b) another small percentage of the original data stored under a different encryption scheme.

Users will then generate queries against the additional piece of data and analyze their results to obtain integrity assurance. Cloud computing, more than any other form of digital communication technology, has created a need to ensure that protection is applied at the inception of the information, in a content

centric manner, ensuring that a security policy becomes an integral part of that data throughout its life cycle.

Summary and Conclusion

Cloud computing has acquired considerable attention from both industry and academia in recent years. Among all the major building blocks of cloud computing, data storage plays a very important role. As mentioned earlier that the most security issues that arise for enterprise through the use of cloud computing is due to the fact of lack of control on the physical infrastructure [6][7]. Enterprises do not know where their data is resided and which security mechanism is applied to protect it. Users require security and privacy to access their personal data objects. Users require secure access to the data for discovery, browsing and computing. In our vision, we propose the use of trusted computing and use of cryptographic computation protocols which supports a dual encryption approach to sensitive data prior to being uploaded to the data cloud storage. To avoid unauthorized access to the sensitive data, any application running in the cloud should not be allowed to directly decrypt the data.

Consequently we surveyed a lot of threats that is associated with cloud data centres. It is anticipated that the approach suggested in this article will contribute to paving the way for securing and building confidentiality in distributed data storage environment within cloud computing platform.

.

Referencing

[1] Boneh.B., D.Crescenz, G., Ostrovsky, R., and Persiano, G. Public key encryption with keyword search in EUROCRPT.2004.

[2] Debora Di Gia Como and Tino Brunzel; Evaluating Cloud Computing-How it differs from traditional I.T outsourcing. May 2010.

[3] EMC. Information-Centric Security. http://www.idc.ptt/resources/ppTs/2007/IT&Internet_security/12.EMC.pdf.

[4] F. John Krautheim, Dhananjay S, Phatak, and Alan T. Sherman. "Private Virtual Infrastructure a model for Trust in cloud computing". TR-CS-10-04, University of Maryland Baltimore country Baltimore, MD,2010; http://www.cisa,umbc.edu/papers/Krautheimtr-cs-10-04-pdf.

[5] Guido Kok. "Cloud computing and confidentiality". May 2010, University of Twente.

[6] Rajikumar Buyya et al. "Cloud computing principles and paradigms". A John Wiley and sons .INC. publication.

[7] Rehan Saleem,. Cloud Computing Effects on Enterprises. Lund University. January,2011.

[8] Richard Chow, Philippe Golle, Markus Jakobsson, Elaine Shi, Jessica Staddon. PARC. Controlling Data in the Cloud: Outsourcing computation without outsourcing control.

[9] Susan morrow. "Data security in the cloud". A John Wiley and sons.inc.publication.

[10] Spiekermann and Cranor(Spiekermann and Cranor 2009). Personal Privacy.