

Exploring Data Security Issues and Scrutinizing them in Cloud Computing Environment

Er. Amanpreet Kaur¹, Dr. Dheerendra Singh²,

¹Department of Information Technology, CEC, Landran, Mohali

²Department of Computer Science & Engineering, SUSCET, Tangori, Mohali
cecm.cse.akb@gmail.com, professordsingh@gmail.com

Abstract

The concept of cloud computing has reached its endeavors and is one of trending topics in world of technology. It is big new idea that can reshape the information technology (IT) services scenario. It is combination of Internet and centralized network servers forming a mesh called CLOUD. These servers are located on remote computers so that we can access applications directly from distant computer by sitting at our place. Cloud is called as Data Center enclosed with boundary called as Demilitarized Zone (DMZ). This zone stores confidential information. In simple words, we can say that cloud refers to Internet and its infrastructure. Cloud computing is an amalgamation of various technologies like Grid Computing, Service oriented architecture [1] and Ubiquitous computing [2]. Since it utilizes every technology, so security is one of the factors that need to be addressed for securing data from outside world.

The paper explores various cloud data privacy issues after studying literatures and studies. Their solutions are being described and conglomerated in the form of holistic cloud computing framework.

Keywords: Cloud computing, Security, Virtualization and Network models

1. INTRODUCTION

The concept of cloud computing was originated from telephony networking scheme. Before evolution of cloud computing, there used to be virtual networks that perform work of connecting multiple computers connected through base coaxial cable. This task is performed in 1990's. The use of virtual networks eliminated the hardware circuits between producers and consumers. But, it did not provide services to multiple users at same time and failed to maintain interoperability with servers and clients.

It leads to evolution of cloud computing that works in distributive environments with multiple sources of

information. Cloud computing is one of increasing trends in world of technology. It is given as name "dujour" says Gartner's Ben Pring [3].

Cloud Computing concept is gaining popularity continuously but it has some common failed perceptions that are governing roots of cloud organizations as listed below:

(i) Data Recovery and Availability: - All applications are designed by considering some laws or rules that are called as Service level agreements (SLA's). There are teams designed to support data availability at anytime. These teams perform following tasks:

- Data Replication
- System monitoring
- Maintenance
- Recovery from failure

(ii) Management Capabilities: - Although there are many multiple cloud providers, but management scale is not satisfactory. There is great need to improve on scalability and balancing features.

(iii) Lack of customer background checks – It is seen that most cloud providers trust unknown users easily and provides them link to register on their sites as a guest. It leads to illegal accessing of customer credentials like credit card and email details. Fictitious accounts can let attackers perform any malicious activity without being identified [4].

(iv) Unawareness about security threats – Mostly, users are not aware of the informative practices and security threats that can govern systems either directly or indirectly. It occurs in most of companies that is due to negligence and careless behavior of clients towards service providers [5].

The paper is organized as follows: Section 2 deals with overview of cloud computing and its deployment or network models. It also puts some light on concept of virtualization to maintain flow of paper. Section 3 lists various issues that are acting as hindrances in privacy of data. It also devised framework incorporating solutions of these issues. Section 4 concludes the given paper and states

future work that can be done to improve process of security.

2. OVERVIEW OF CLOUD COMPUTING

Cloud computing is massive information technology that helps in performing large computational tasks in easier way. It provides services in seamless manner which means that users can use services according to their needs and pay for only those services that they use. Cloud computing deals with relocatable services, provides data for cost and centralized database of servers. Relocatable services mean that services can be provided to remote desktops on remote locations if data is available in cloud. So, it is very important that data must be secure and present in cloud before accessing it.

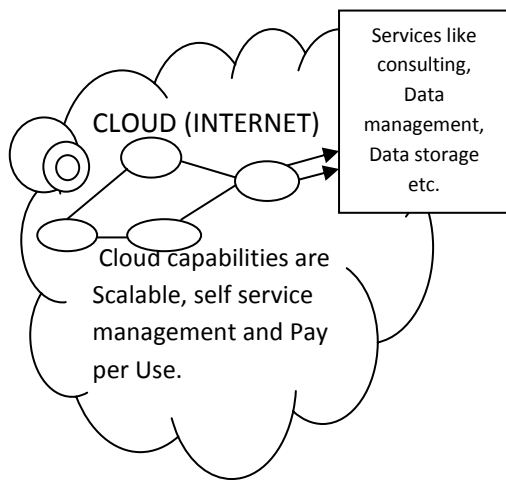


Fig 1: Capabilities of Cloud

2.1 Deployment Models

They are also referred to as cloud computing types- Public Cloud, Private Cloud and Hybrid Cloud. *Public Clouds*: - They provide services to anyone globally with an Internet connection. They are owned by company which has its certain services. E.g. YouTube and yahoo are sites from which user access data in form of emails, attachments, videos from any device that has internet connection.

Private Clouds: - They provide services privately and has private network to limited number of people using that network.

Hybrid Cloud: - It is combination of both public and private cloud models. It has ability to add private cloud with resources of public cloud that helps to manage unexpected workload on cloud.

S.No.	Public Cloud	Private Cloud	Hybrid Cloud
-------	--------------	---------------	--------------

1.	Simple and easy to use.	Monitoring is needed to control latest software updates	Most efficient (combination of both)
2.	Widely accessible.	limited accessible	Used to reduce work load.
3.	Less costly and reliable	More costly and less reliable	Most costly and most reliable.
4.	Suitable for handling large workload pressure	Not suitable for large workload pressure	Suitable for handling large workload pressure
5.	No space allocated for data center	Largest space allocated for data center.	Average space allocated for data center

Table 1: Public vs. Private vs. Hybrid models

2.2. Virtualization in Cloud Environment

Why there is need of virtualization?

The existing issues are capable for performing computations in Local area networks (LAN's) environments. These environments have fixed bandwidth, less availability of bandwidth index, static allocation of resources and non-distributed data centers for handling customer requests.

So, there is need to extend this service of cloud computing to wide area networks (WAN's) that possess dynamic and scalable allocation of resources, high bandwidth index and large distributed databases.

What is virtualization?

Virtualization is defined as treating a single physical resource like operating system, storage devices to act as multiple virtual servers. It is treated as like black box. Users are concerned about using these services instead of knowing how they originate into cloud.

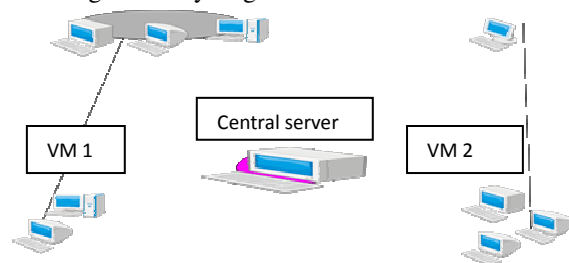


Figure 2: Virtualization process

3. DATA SECURITY ISSUES

Since data and services are provided by vendors to users through wide range of networks on Internet. It is obvious

that provisioning of too many services on Internet will lead to some privacy concerns [6]. They are described as:

(a) Data Access: - There is need to maintain some cloud standard or security measure to prevent threat of accessing sensitive information from cloud. Hackers and malicious intruders are active always and they can hack confidential data to use it in other ways. The existing cloud system is not viable for deploying cloud models because different areas in cloud are associated with different risks.

(b) Lack of trust and authentication: - Various cloud providers are providing their services at “pay per use” package. But they have forgotten to maintain trust between user and themselves. They does not provide any audit mechanism for monitoring services and transactions executed by different users on different cloud platforms.

The question remains still in user mind “How user can trust that his data is safe”? Lack of trust is one of main factors that has created feeling of fakeness and stopped various organizations to adopt cloud standards easily. Although cloud is best platform for storage of too many services but without its security, it is of no use. Cloud computing services must be improved in legal protection also [7].

(c) Data Segregation: - The word segregate means to isolate from group of people or platforms. Multi-tenancy is one of major characteristics in cloud computing which states that multiple users can store their data at multiple locations and can access services of different layers. This led hackers to inject into client code of other systems and can steal their data. There must be some boundary between different layers so that service should separate data from different users in secured way.

(d) Data Breaching: - The word breach means violation or infraction. In cloud computing context, it means data lying on different platforms in cloud environment may cause violation of policies and service legal agreements. There must be some strict SLA’s that needs to be employed in order to prevent violation of data.

(e) Virtual Machine (VM) Security:- Performing different instances of same task on single physical machine is major task of virtualization.

Since virtualization is appreciable phenomenon but it also governs with monitoring and metering issues. Current virtual machine monitors (VMM) are not able to isolate instances from each other because anyone can login as guest and access data in cloud.

(f) VM Stall: - Most companies are working on virtualized servers. Large number of virtual machines is increasing day by day due to virtualized deployment. VM Stall occurs due

to lack of trust in administrators and performance management.

(g) VM Sprawl: - Easy provision of virtual machines may lead to out of control of pop-up blockers on various cloud sites. The blockers will allow ads on sites simultaneously which leads to lack of management and handling of client requests.

(h) Cloud Standards: - They are needed to maintain interoperability and security among various cloud providers like Google, Amazon etc so that user cannot leave its respective providers and migrate to others. It is duty of statutory bodies to look into privacy agreements and establish some common standards for all cloud providers.

(i) Excessive bandwidth problem:- Utilization of services varies from user to user depending on their capacity of accessing services. It may be possible that single user depending on its service may consume more than share of useful bandwidth. It is serious problem as it leads to wastage of resources.

(j) Verification of Identity and signatures:- ID management is associated with every individual working in cloud environment. It is one of issue that is still not adopted by various companies. The cloud providers have right to check that who is utilizing their services and in what way. In such scenarios, provisioning of users is of great concern.

(k) Web application security:- Web applications at network level pose serious threats by allowing users to manage their activities from central locations rather than at individual customer’s sites. The applications are distributed over network so they can be accessed from their remote locations also.

3.1. Proposed Data Privacy Framework

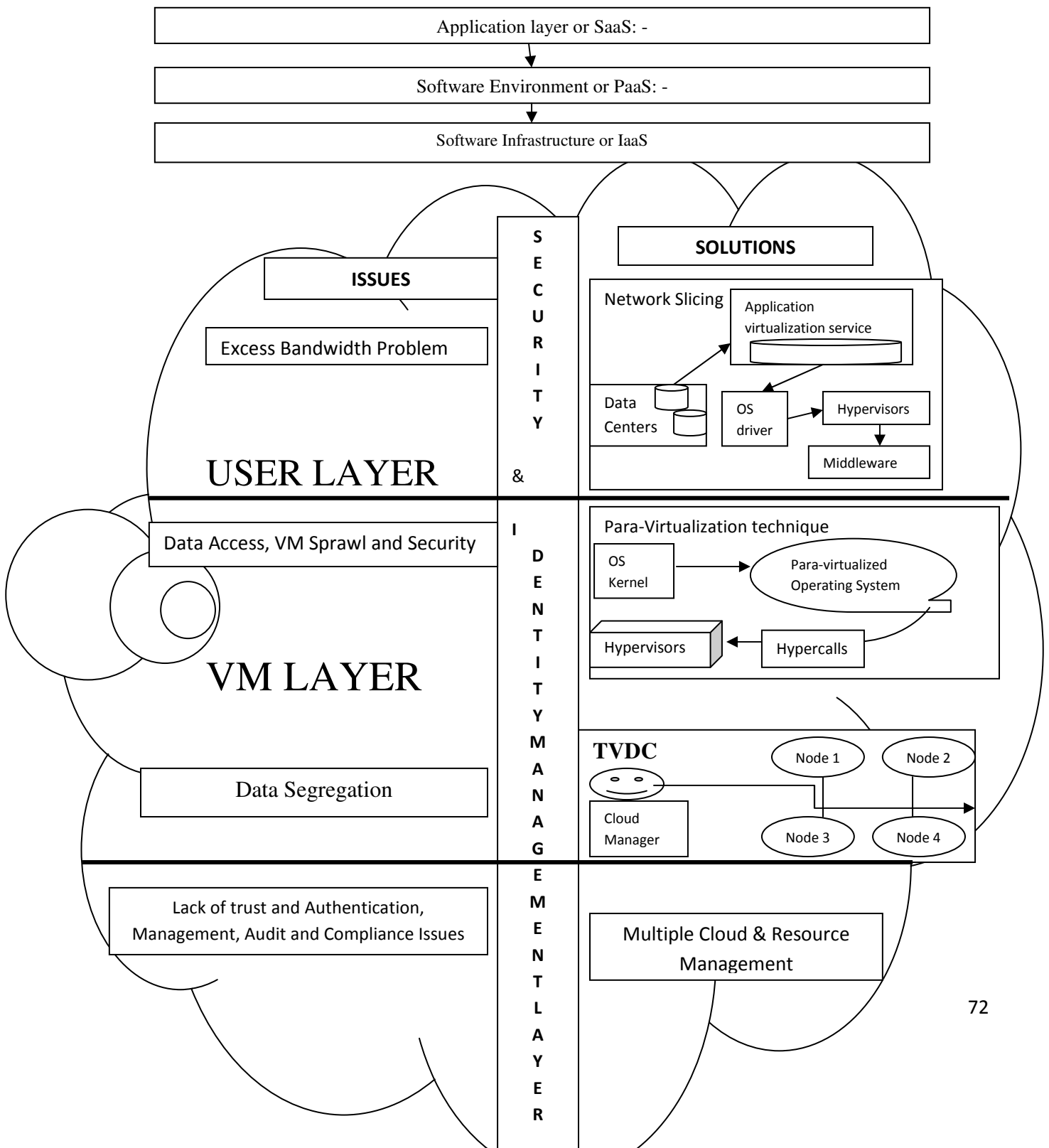
The model is designed in accordance with aforementioned privacy issues that need to be addressed in public cloud environment. These issues have become threats to cloud environment and thus cloud computing architecture must be scrutinized with their respective solutions.

(a) Continuous Research by few statutory bodies

Various research studies and experiments are going on to maintain trusted cloud environment. Several groups and organizations are finding solutions to cope with security threats. Few of them are Cloud Security Alliance (CSA)[8], ITU Cloud Computing focus group [9] and IEEE Cloud computing standard study group [10] gathered various solutions and maintain coordination with other forums for information assurance in cloud.

(b) Solutions provided by researchers

W. Tsai [11] proposed four tier architecture that includes process of web development and provides separate security layer. Berre [12] proposed framework with use of X10 language for achieving concurrency in cloud environment. Singh et.al [13] designed resource isolation approach to ensure security of data.



3.2. Data Privacy Solutions

(c) Solutions defined in above model

- Network Slicing:- In cloud computing, it is possible that single user can access resources more than its capacity depending on their services. It leads to wastage of resources and bandwidth. This technique allocates set up of VM's among each user in order to slice network for bandwidth conservation. Hypervisors act as interface between OS drivers and hardware. The resources are passed from one VM to other machines. Then OS driver communicates with hardware to perform given task.
- Trusted Virtual Data Center Implementation (TVDC):- It is solution to data segregation issue that says that resources must not get leaked to other customer area. So, TVDC uses isolation policy that separates hardware resources and users workload of data. This isolation policy includes major tasks viz. label (for identifying VM's which are associated with users) and allows VM's to run on same TVD.
Label has cloud manager (CM) that responds to set of nodes arranged in cluster under control of hypervisors.
- Para-virtualization involves modifying the OS kernel to replace non-virtualizable instructions with hyper calls that communicate directly with the virtualization layer hypervisor.

4. CONCLUSIONS AND FUTURE WORK

The paper studies various data privacy issues that are governing cloud computing technology. The issues include data breaching, segregation of data, lack of trust and authentication and many more. The concept of virtualization has been introduced in order to access data from cloud in faster way. Virtualization has challenges like VM Sprawl, VM Escape but their subsidiary solutions are provided in given paper.

A data privacy framework has been devised that incorporates solutions of data privacy issues in holistic way. Para-virtualization technique is being used to deal with data authentication and virtualization challenges. It also includes concept of trusted virtual data center (TVDC) to maintain isolation of resources.

Future Works

There can be some statistical model designed that predicts occurrence of attacks by intruders and hackers by involving some equations. TVDC model can be set up and implemented using some privacy model or repudiation model.

REFERENCES

- [1]. L-J Zhang, EIC Editorial: Introduction to the Body of knowledge areas of services computing IEEE Transactions on Services Computing, 1 (2), pp 62-74, 2008.
- [2]. L. Klein, A Vision for the Internet, ST Journal of Research, 2 (1), Nov 2005, pp 4-5.
- [3]. G. GRUMAN, E. KNORR, What cloud computing really means. *InfoWorld*, (2009, May). [Online]. Available: <http://www.infoworld.com/d/cloudcomputing/what-cloud-computing-reallymeans-031>
- [4]. Heinrich, M., Grüneberger, F.J., Springer, T., Gaedke, M.: Exploiting annotations for the rapid development of collaborative web applications. In: WWW, pp. 551-560 (2013)
- [5]. Cloud Security Alliance (2012) Top Threats to Cloud Computing V1.0. Available: <https://cloudsecurityalliance.org/research/top-threats>
- [6]. Cloud Security Alliance: <http://www.cloudsecurityalliance.org>
- [7]. Blaze M, Feigenbaum J, Ioannidis J, Keromytis AD. The role of trust management in distributed systems security, secure Internet programming, issues for mobile and distributed objects. Berlin: Springer-Verlag; 1999.p.185-210
- [8]. Cloud Security Alliance. Security best practices for cloud computing, 2010b <<http://www.cloudsecurityalliance.org>>
- [9]. ITU, 2013. Cloud Computing Focus Group. <<http://www.itu.int/en/ITU/focusgroups/cloud/Pages/default.aspx>> [Accessed :December 2014]
- [10]. IEEE CCSSG. IEEE Cloud Computing Standard Study Group.

<<http://www.computer.org/portal/web/sab/cloud>>

[Accessed :December 2014]

- [11]. Tsai W, Jin Z, Bai X. Internetware computing: issues and perspective. In: Proceedings of the first Asia-Pacific symposium on Internetware. Beijing, China: ACM; 2009. p. 1–10.
- [12]. Berre AJ, Roman D, Landre E, Heuvel WVD, Skar LA, Udnaes M, et al. Towards best practices in designing for the cloud. In: Proceedings of the 24th ACM
- [13]. Raj H, Nathuji R, Singh A, England P. Resource management for isolation enhanced cloud services. In: Proceedings of the 2009 ACM workshop on cloud computing security, Chicago, Illinois, USA, 2009, p. 77- 84.
- [14]. RaoMikkilineni, Vijay Sarathy "Cloud Computing and Lessons from the Past", Proceedings of IEEE WETICE 2009, First International Workshop on Collaboration & Cloud Computing, June 2009
- [15]. Quiroz A, Kim H, Parashar M, Gnanasambandam N, Sharma N, 2009, "Towards workload provisioning for enterprise grids and clouds", IEEE/ACM international conference on grid computing, pp 50-57