Secured Data Migration from Enterprise to Cloud Storage – Analytical Survey

Neetu Kishore[1] and Seema Sharma[2]

Submitted in November, 2015; Accepted in January, 2016

Abstract - Everything in Cloud is an emerging concept, with work in multiple areas right from the Infrastructure, to middleware, to applications and to data security. Every domain has its own flavor of Cloud. Technologies and techniques for Cloud has come a long way and is still evolving. Organizations are fast recognizing the value of migrating to Cloud but the biggest concern is around data security. The migration of sensitive data to a public cloud domain has risks associated with data loss, information theft, confidentiality and other vulnerabilities. There are security measures deployed at multiple points but the question of secured end to end data transmission still remains unanswered. Objective is to understand the security parameters that can help build a framework for secured data transmission. With reference to many research papers and reviews, the analysis is to focus on the security of data in transit. Pragmatic observation is that the data at rest in the Enterprise and the Cloud have fairly good security measures, but the data in transit is vulnerable. There is a future scope to build a security framework for the transit data. This paper will highlight few of the security aspects of Cloud that have been developed and the gap around the security of data being migrated from Enterprise storage to Cloud Storage.

Index Terms – Availability, Cloud, Cyber, Confidentiality, cryptography, encryption, Homomorphism, Integrity, vulnerability.

NOMENCLATURE

IaaS-Infrastructure-as-a-service; PaaS- platform-as-a-service; SaaS- software-as-a-service; CSP- Cloud Service Provider; SSL – Secured Socket Layer, TLS- ; HTTP- Hyper Text Transfer Protocol; VPN- Virtual Private Network; DSL- domain-specific language; EDSL- embedded domain specific language; IP- Intellectual Property; OS- Operating System; DES- Data Encryption Standard; CSA- Cloud Security Alliance.

1.0 INTRODUCTION

Cloud is the new paradigm of traditional data centers with enhanced features to offer services in a shared environment on virtualized systems, capability of multi-tenacity, scalability and pay-per-usage pricing model. Organizations which have non-IT as core business, started looking at these IT services, moving away from their infrastructure investments and IT asset ownership.

Neetu Kishore, M.Tech Student at BIT Mesra, Noida Campus, India. Email – neetukishore31@gmail.com
Seema Sharma, Asst. Professor CSE, BIT Mesra, Noida Campus, India. Email – seema@bitmesra.ac.in

Typical Cloud services started with moving out the Infrastructure layer from in-house to a shared data center environment, followed by platform and now applications. The main objective of cloud computing is to make better use of distributed resources and solve large scale computational problems. [1]

While standardization of the cloud computing is still in progress, an agreed framework for describing cloud computing services has been accepted commonly as "SPI." This acronym stands for the three major services provided through the cloud: software-as-a-service (*S*aaS), platform-as-a-service (*P*aaS), and infrastructure-as-a-service (*I*aaS)" [2].

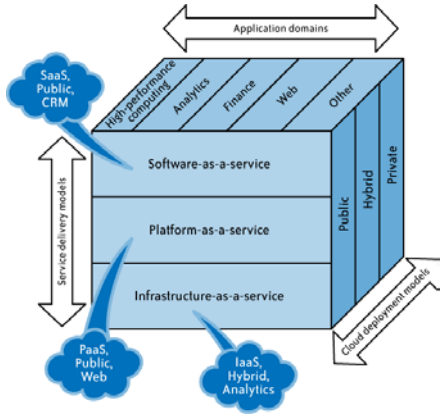

Figure1.1 SPI Service Model illustrating the relationship between services, uses, and types of clouds [2]

This framework has helped Cloud computing services get structured in layers of service offering by the CSPs (Cloud Service providers) [3]. The CSP maturity evolved and from the standard services, they started offering customized and value added services in multi-tenant segregation on shared cloud platform. Moving from monolithic onsite data centers, the virtual servers through browser access has provided a substantial throughput and maintenance in the software and hardware environment. [4]

2.0 SECURITY IN CLOUD

In the existing Cloud environment, two areas play critical role for sustainability – Data Security and Standardization. These are still at an early stage and have major scope in future for research, considering the complexity of the Cloud computing architecture. Cloud Security and related research reviews have shown that the area of concern emerged as security of transit data from enterprise to cloud storage. There has been many researches done on how to keep data secured at rest either in the Enterprise or in Cloud, but Organizations are paranoid

about security of data during transmission. There are techniques applied for secured data migration for specific requirement but there is a lack of framework to address all types of data migration in a secured way from a private environment to a public domain of cloud.

2.1 Security concerns

In the existing Cloud environment, two areas play critical role for sustainability – Data Security and Standardization. These are still at an early stage and have major scope in future for research, considering the complexity of the Cloud computing architecture. Another important area to understand in the data security is the types of risks to data for the cloud migration. As per Cyber Security literature, there are three threat classes in cyber-attacks – piracy, tampering and reverse engineering. [5]
Piracy – This is about the copying of data to unauthorized locations. Tampering – Malicious manipulations of computing logics would impact the businesses. Reverse Engineering – The vulnerabilities in the applications and system software are exploited to steal the Intellectual Property (IP).

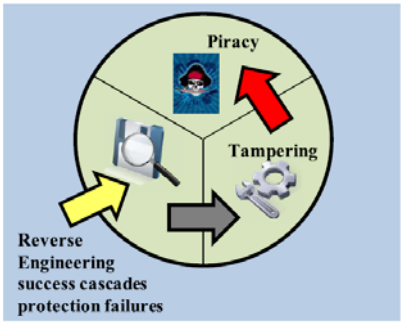

Figure 2.1 Graphical depiction of attack on a computing asset [5]

With increasing threats [6], security measured have been applied at various levels, hardware, OS, applications and other protection tools based on hardware or software programming. No one system can counter various attacks and to put all measures together, the impact is on the cost and performance of the system. Organizations worked on criticality of the data and arrived at an optimal solution, understanding the associated risks. Only hardware was sufficient to guard the outer layer and prevent external attacks, but the other systems still remained vulnerable.

2.2 Secured data transmission

The greatest fear of organizations to migrate their Enterprise data to cloud storage is loss of control. Every organization wants to have their data privacy and security at all levels to ensure Confidentiality, Integrity and Availability of Enterprise data. Cyber threats have created the fear that data is highly vulnerable to multiple types of risks during transmission to cloud. Cloud Service Providers have matured to build security models on their hardware and software but the threat at the data access and migration level still remains vulnerable to attacks.

Protecting data in motion is a major concern. [7] Primary method of securing data from exposure of network media is encryption, and it could be applied in two main ways: by encrypting the data itself to protect it or by protecting the entire connection. SSL and TLS are often used for web traffic with HTTP (Hyper Text Transfer Protocol). When SSL/TLS are used, the connection between the two systems over the network is encrypted and is very specific to an application or a protocol. The second approach is to encrypt all the traffic through a VPN (Virtual Private Network) connection.

3.0 OBSERVATION ANALYSIS

For the security of data in the cloud computing, there is a need to build a secured framework for the migration of data. The framework should provide guidelines to ensure there is end to end security of data from enterprise to cloud. Additionally, there are technical Issues concerning the data security. There are three parts to be considered – the security at storage level, the security at computation level and the security at migration level. Data from a secured Enterprise computing environment migrated to public cloud has to be ensured for retaining its confidentiality and integrity and once stored in cloud storage has to guarantee the availability without any loss.

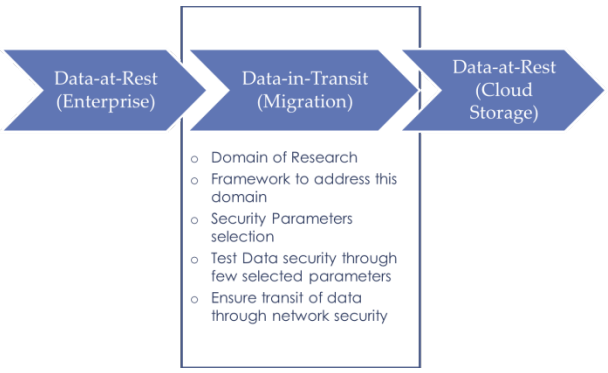

Figure 3.1Conceptual model for the framework

The security measures at different levels ensure Security at either periphery, or within system or at applications level. For a comprehensive secured data handling in cloud should have a structure with relevant parameters. A conceptual model in Fig 3.1 would be a starting point of putting the pieces of the Security framework. Through survey analysis, itwas observed that only programming techniques or security at software level cannot protect the data when it moves from one physical environment to another and resides on a shared platform. Any outgoing or incoming data with application based security only could be compromised. This security mechanism could be defeated by attack tools that intercept calls from the software to the operating system (OS) and falsified or altered the outgoing or incoming data [8].
In trusted computing and computation-supporting encryption, life in the cloud can be advantageous from a business intelligence standpoint over the isolated alternative that is more common today. [9]

4.0 SECURITY TECHNIQUES

Cryptography [10] is the science and study of Secret writing. In cloud computing the security concepts are to be understood in depth for relevant application and modification of techniques. Any information that is transmitted over electronic wire is vulnerable in two ways – passive wire-tapping and active wire-tapping. Passive does not do any detection, just intercepts the messages where as Active tampers and modifies the data on transit. In the emerging digital world, while secrecy of data is required, authenticity is equally important. DES standards could help in secrecy but cannot guarantee the authenticity.

The Encryption Algorithms use different types of ciphers. Brief explanation of few of the key cipher techniques is as follows. [10] Transposition ciphers rearrange characters according to some scheme. Substitution ciphers are of four types – simple substitution ciphers which replace plaintext with a cypher text with a one to one mapping, Homophonic substitution ciphers are similar but with a one to many mapping, Polyalphabetic substitution cyphers use multiple mapping of plaintext to cipher text and Polygram substitution cyphers allow arbitrary substitution of group of characters.

Vulnerability is the procedural weakness in the hardware or software that allows the attackers to gain access to networks and penetrate into any system inside the network. [11] The security analysis for leakages of data and prevention through various means is very critical for the selection of the parameters.

In the data transmission method, transferred data is encrypted in the upper-layer on top of the transport layer instead of using IPSec or SSL.[12] Without doing any modification on IP layer, encryption can be pre-processed at upper layers. This will improve performance.

5.0 SELECTED PAPER SYNOPSIS

The review details for selected research papers to understand the security of data in the cloud computing environment, common trend has been around securing data at rest, either in the enterprise or in cloud storage. Very limited work has been so far done on the security of data in transit. Different security mechanisms and also vulnerabilities have evolved as the cloud computing trend has progressed.

Using homomorphic encryption [13] and secure multiparty computation, cloud servers may perform regularly structured computation on encrypted data, without access to decryption keys. The core of domain-specific language (DSL) is implemented as a Haskell library, an embedded domain specific language (EDSL). The implementation includes Shamir secret sharing and fully homomorphic encryption; both use SSL network communication between clients and any number of servers. While homomorphic encryption and secure multiparty computation are based on different cryptographic insights and constructions, there is a surprising structural similarity between them. This similarity is also shared by so-called partially homomorphic encryption, in which the homomorphism property holds only for certain operations [14]. In cryptographic domain, the computation on encrypted data needs to be worked without any change in the expected result. The challenge is in the data being transmitted securely through the public transmission media.

One of the emerging trends in cloud storage is the repository of data being captured during web transactions and used for analytics to provide valuable information to the business. With this trend, there is sensitive data that also flows through public network and the concerns are around securing this type of data.

Big data has the fear of sensitive information getting leaked out. The mechanisms to protect this data are still not matured for the shared platforms. Big data is an advanced form of storage in Cloud which allows the transaction of data and sharing of information. Research work still undergoing for a complete secured environment for safe data movement. [15]This paper proposes a framework for secure sensitive data sharing on a big data platform, including secure data delivery, storage, usage, and destruction on a semi-trusted big data sharing platform. A proxy re-encryption algorithm based on heterogeneous cipher text transformation and a user process protection method based on a virtual machine monitor, which provides support for the realization of system functions. The framework protects the security of users' sensitive data effectively and shares these data safely. At the same time, data owners retain complete control of their own data in a sound environment for modern Internet information security.

To understand the Security vulnerabilities [16] for data moving across the public and private networks, it is important to analyze the various possible attacks through internal or external sources. There is always a concern on the confidentiality, Integrity and Availability of data where there are security concerns and hence, the best way to observe these risks is through threat attack simulations.

6.0 SURVEY COMPARISION

Some of the factors compared during the survey are as summarized in the table below:

Sr. No.	Factors Compared	Observations
1	What needs security?	Company Data, personal information,
2	Threat parameters	Vulnerability, confidentiality, integrity
3	Shared resources	Concern of data leakage
4	Data in transit	Greatest risk
5	Key security parameters	Network protocols, secured access, cryptography,

7.0 CONCLUSION AND FUTURE SCOPE

The reviews of the research papers on cloud related security show that end to end security in cloud data storage is a challenge due to complexity of the environment. While there are different security measures for data protection in the common computing environment, the Cloud architecture needs new techniques for security. It is important to realize that from

a scientific standpoint, there is no absolute notion of security. [17]

In Cloud computing, data migration from and to the enterprises is very critical. Data constantly needs to be moved from one place to the other either for computation, storage or simple access. This needs a better understanding of the encryption techniques for data at rest and data in motion. Encrypted data, for access at multiple stages, need to ensure that the integrity is maintained throughout. There are multiple active adversaries, against which data has to be guarded while in motion.[18, 19] Cost vs Efficiency need to be well understood for end to end cloud security considering the technology and regulatory challenges. For a higher security, the cost implication is going to be high and also the data at every stage, being run for a security procedure, will impact the efficiency of the system. There is a lack of Framework for data in motion and security against active adversaries has scope of future work.

Cloud Security Alliance (CSA) [20] is gathering solution providers, non-profits and individuals to enter into discussion about the current and future best practices for information assurance in the cloud. The CSA has identified thirteen domains of concerns on cloud computing security [21].

8.0 ACKNOWLEDGEMENT

I would like to thank my coach and guide, Ms Seema Sharma for being a great supporter of the research topic on Cloud and this survey review paper. I thank Mr. Santanoo Pattnaik, visiting faculty of BIT who has been very helpful on my research topic. I would also like to thank Mr. Chirag Shah, Technical specialist on Cloud technology at General Electric.

REFERENCES

[1]. Matthew N. O. Sadiku, Sarhan M. Musa, and Omonova D. Movo, "Cloud Computing: Opportunities and Challenges", IEEE Potentails, pp. 34, 7 January 2014.

[2]. Tim Mather, Subra Kumaraswamy and ShahedLatif, - Cloud Security and Privacy -published by O'Reilly Media, Inc. Published date - September 28, 2009.

[3]. Bansi Khimani, and Kuntal Patel, "A Novel Model for Security and Data Access for Jointly Accessing the Cloud Service", BIJIT-BVICAM's International Journal of Information Technology. January - June, 2015; Vol. 7 No. 1.

[4]. "Addressing Data Security Challenges in the Cloud" – A Trend Micro White Paper, pp.2, July 2010.

[5]. Kelce S Wilson and MugeAyseKiy, "Some Fundamental Cybersecurity concepts", IEEE Access, Volume 2, pp. 117-118, February 2014.

[6]. Bilal Maqbool Beigh, "Framework for Choosing Best Intrusion Detection System", BIJIT-BVICAM's International Journal of Information Technology. January - June, 2015; Vol. 7 No. 1.

[7]. Jason Andress - The Basics of Information Security - Understanding the fundamentals of InfoSec in theory and practice; Second Edition; Syngress Publication, 2014.

[8]. Z. Wang, X. Jiang, W. Cui, and P. Ning, ``Countering kernel rootkits with lightweight hook protection," in Proc. 16th ACM Conf. Comput. Commun. Security, Nov. 2009, pp. 545_554.

[9]. Richard Chow, Philippe Golle, Markus Jakobsson, RyusukeMasuoka, Jesus Molina, Elaine Shi, Jessica Staddon, "Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control", CCSW'09, November 13, 2009, Chicago

[10]. Dorothy Elizabeth Robbling Denning - Cryptography and Data security; Addison-Wesley Publishing company, 1982.

[11]. Chris Sanders and Jason Smith and Technical Editor David J. Bianco - Applied Network Security Monitoring, collection, detection and analysis; Elsevier Inc, 2014.

[12]. Pradnyesh Bhisikar and Prof. Amit Sahu**, "**Security in Data Storage and Transmission in Cloud Computing", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, pp. 5, March 2013

[13]. C. Gentry, "Fully homomorphic encryption using ideal lattices," in STOC, 2009, pp. 169–178.

[14]. John C. Mitchell, Rahul Sharma, Deian Stefan, and Joe Zimmerman,"Information-flow control for programming on encrypted data", IEEE 25th Computer Security Foundations Symposium, 2012.

[15]. Xinhua Dong, Ruixuan Li, Heng He, Wanwan Zhou, ZhengyuanXue, and Hao Wu, "Secure Sensitive Data Sharing on a Big Data Platform', TSINGHUA SCIENCE AND TECHNOLOGY.Volume 20, Number 1, February 2015.

[16]. Neha Mishra, ShahidSiddiqui, and Jitesh P. Tripathi, "A Compendium Over Cloud Computing Cryptographic Algorithms and Security Issues", BIJIT-BVICAM's International Journal of Information Technology. January - June, 2015; Vol. 7 No. 1.

[17]. Carl Landwehr, Dan Boneh, John C. Mitchell, Steven M. Bellovin, Susan Landau, and Michael E. Lesk, "Privacy and Cybersecurity: The Next 100 Years", Proceedings of the IEEE, Vol 100, May 2012.

[18]. Meenu Dave, Mikku Dave, and Y. S. Shishodia, "Cloud Computing and Knowledge Management as a Service: A Collaborative Approach to Harness and Manage the Plethora of Knowledge", BIJIT-BVICAM's International Journal of Information Technology. July-December, 2013; Vol. 5 No. 2.

[19]. Vaibhav Rana, "Innovative Use of Cloud Computing in Smart Phone Technology", BIJIT-BVICAM's International Journal of Information Technology. July-December, 2013; Vol. 5 No. 2.

[20]. CSA http://www.cloudsecurityalliance.org.

[21]. Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing, V2.1, http://www.cloudsecurityalliance.org/guidance/csaguide.v2.1.pdf.