Data Security in Cloud Computing: Major Concerns and Implications

Prof Shubhangi A Shinde
Head,Dept of MCA,
G.H. Raisoni College of Engg& Mgt,
Wagholi, Pune, India
shinde.shubhangi@gmail.com

Dr. Vinay Chavan
Head, Dept of Science
S.K. Porwal College,
Kamptee , Nagpur
drvinaychavan@gmail.com

Abstract:**Many organizations are realizing benefits of cloud computing by putting their data in cloud. Cloud adoption may lead to efficiency and effectiveness in developing and deploying the applications which leads to cost benefit. According to the recent study the main obstacles for adopting cloud computing are security, privacy and legal matters as it provides marginal solutions for adoption of cloud. The main challenges faced by cloud adoption are data loss, data leakage, data integrity and data availabilityUsers of the cloud want to ensure that data in the cloud are subject to unauthorized access or corruption or modification and disclosure.In this paper we have discussed data security challenges in cloud computing and implications associated with it.**

Keywords**: cloud computing, data security issues in cloud computing, data integrity and confidentiality, benefits of cloud computing, privacy and data protection in cloud.**

I INTRODUCTION

Cloud computing is gaining popularity as it provides variety of different type of services and infrastructure models which will have its own advantages and disadvantages[1]. Cloud model provides various computing based services to users, so they do not require developing expertise. For example Amazon's EC2 service provides clients scalable servers and S3 provides scalable storage to clients. Although cloud computing provides many potential advantages from consumers perspective, security concerns are major barriers for adoption of cloud. According to survey conducted by Gartner [15] 70% of CTO's believed data security and privacy are major concerns for organizations not using clouds. Amazon's storage service was interrupted twice in February and July 2009.

Cloud adoption will have its own positive and negative effects on the data security of consumers. According to analysis by 2012 size of cloud computing will reach upto $60 billion to $ 80 billion 10% of global IT market [1]. According to [2] in cloud computing the provider's computing resources are pooled to serve multiple consumers using a multitenantmodel, with different physical and virtual resources dynamically assigned and reassignedaccording to consumer demand. Information and personal data are transferred from one location to another location and customer does not have knowledge about the exact location of the data. It is virtually not possible for the data owner to confirm security conditions of all the server locations that might be used to house the data. In cloud computing environment data and application is controlled by service provider which leads to its safety concerns from internal as well as external sources.Users are at the mercy of their service provider for accessibility and varcity of their data.

ENISA[3] has pointed main concerns related to cloud model are data protection and data security, confidentiality of information and intellectual property, vendor lock in, law enforcement access, CSP professional negligence, subcontracting of cloud services and CSP change of control. Despite security concerns in cloud computing environments, advantages such as elastic-scale, pay as you go, reduced maintenance cost and flexible infrastructure platform are compelling reasons for enterprises to attract towards cloud computing environments.

Orgnaization of this paper will be as follows. Section II will describe Security benefits of Cloud computing for data storage. Section III will describe Security concerns of cloud computing. Section IV will describe implications associated with security concerns and Section V will conclude the paper.

II SECURITY BENEFITS OF CLOUD COMPUTING FOR DATA STORAGE

i. Built in back up and safeguards that is not found in desktop and enterprise servers.
ii. Irrespective of size of the industry all the users receive same level of security.
iii. Security patches and updates are deployed immediately for all the users in cloud model unlike desktop servers.
iv. The technological expertise provided by cloud providers cannot match by in-house operating systems used by the corporate.
v. Large volumes of data in cloud can be cheap solution for disaster recovery and data storage.
vi. Cloud providers try to use state-of-the art intrusion detection systems and on-demand security control.
vii. Virtual servers, data storage, network capabilities are load balancing and automatically expandable. Resources are allocated as needed and loads can be transferred automatically to better locations, producing a robust, reliable service.

III DATA SECURITY CONCERNS IN CLOUD COMPUTING

Although huge data centers are established in cloud, data storage in cloud is not trustworthy. Cloud has a security challenges like data integrity, data lock in, data remanence, data provenance, data confidentiality, and user specific privacy concerns. Migration to cloud model will have its implications on applications and services [5]. According to [5] Data loss or leakage can have severe impact on business, brand and reputation, employee, partner, and customer morale and trust. An example of this is in 2009 Spamhaus, a real time blacklist provider listed all Amazom EC2 IP addresses on its blacklist spamming. It resulted into most network devices automatically rejecting email emanating from anyone using an EC2 IP address. Loss of core intellectual property could have competitive and financialimplications also. Depending upon the data that is lost or leaked, there might be compliance violations
and legal ramifications. Major data security concerns in cloud computing are

C1: Data Segegregation and protection
C2: Data Integrity and Privacy
C3: Data Ownership
C4: Data Location
C5: Data Remanence
C6: Auditing of Service provider
C7: Threat and Vulnerability Management
C8: Encryption Key Management
C9: Support for Business Continuity Process & Disaster Recovery.
C10: Data Provenance.

IV IMPLICATIONS ON SECURITY CONCERNS

In this section we will discusstechnological implications of security concerns and related research in the field.

i. *Implication C1*: In multi-tenant world of cloud computing virtualized servers exist in one physical box. In such situation it is possible to place an attack on virtual server on the same hardware as target virtual server and then build side channel between two that could enable SSH keystroke timing attack to learn passwords. In this environment it is difficult to identify malicious activity conducted by user which could include “brute force” attacks, DOS attacks, scanning of co-tenants confidential information. Robust controls are required to be provided by cloud provider isolation of one user’s data from others.
ii. *Implication C2:* Data center related measures like intrusion detection systems, network security, physical access and authentication is required. TIA-942 Data Center Standards describes the standard for data center infrastructure. Different encryption decryption mechanisms can be established. Encryption with key reduces computational speed and cost for user storing data on server. Homomorphic encryption enables ciphertext processed directly without decrypting it. Various Distributed coding schemes like erasure encoding has been studied and used [6]. A scalable privacy preserving algorithm for data outsourcing is proposed and used distribution on multiple data provider sites and information theoretically proven secret sharing algorithms as the basis for privacy preserving outsourcing [8]. NetDB2 directly addresses two of the main challenges in developing databases as a service, namely data privacy and performance[9].
iii. I*mplication C3:* Cloud computing has the technology to potentially generate new information derived from the data made available from users. Information privacy enforcement in cloud environment by withholding data ownership information from cloud nodes is evaluated [18]. Nodes that have access to private data in unencrypted form do not know who owns it, what role their computations play in the larger

computational task, or to whom their computation results are ultimately delivered. To provide this data ownership privacy, the cloud's distributed computing resources are leveraged to implement an anonymizing circuit based on Tor, through which users submit private data and jobs.

iv. *Implication C4*: In cloud model data could be stored on many jurisdictions around the world. Users could face the possibility of the "laws of the land" in which government or other third parties would access the data of user in spite of the what contract is made between the cloud provider and the user. Standards around outsourcing auditing (SAS70) govern cloud based outsourcing vendors. US Federal and other international laws such as the Electronic Communication Privacy Act (ECPA) can govern concerns for data privacy in cloud. Storage of data in cloud requires fulfilling state and federal privacy and data security law requirements. Some US legal and regulatory requires data owners to ensure that their third party service providers are capable of maintaining privacy and security of personal information entrusted to them [1].Federal privacy laws that restrict the activities of service providers are Health Insurance Portability and Accountability Act (HIPAA) and Gramn-Leach-Bliley (GLB) Act.HIPAA of 1996, Pub. L imposes significant restriction on the disclosure of protected health information. Any HIPAA-covered entity would first have to negotiate and enter into a business associate agreement with a cloud provider before it could store records containing PHI in a cloud computing facility. GLB's Privacy and Safeguards Rules restrict financial institutions from disclosing consumers 'nonpublic personal information to non-affiliated third parties

v. *Implication C5*: Any critical data must not be only protected against unauthorized access and distribution, but also securely deleted at the end of its life cycle. Institutions related to health, defense and financial has to ensure no data is left on the disk from where it is exposed to the risk of being recovered by malicious users. It is known as data remanence. When you are having full control of file server, you can overwrite corresponding sectors to overwrite data to destroy physical trace of file. But in cloud environment this solution is not possible because you have not given physical device, but higher level of abstraction like file systems. this issue can be easily solved in encrypted storage scenarios by throwing away the key, secure deletion of unencrypted data in processing scenarios still remains a huge problem which cannot be solved without the help of the CSP. One solution can be a third party auditor who offers this service to user. Trusted Platform Module (TPM) can provide hardware-based verification of hypervisor and integrity of the virtual instance running. With the help of such methods, trusted log files and trusted deletion of data could be theoretically provided to the customer.

vi. *Implication C6*: A third party data auditing would provide guaranteed and unbiased auditing result for storage auditing in cloud computing. A third party auditor can have more expertise and capabilities and can do a more efficient work and can convince both cloud service provider and owners[14]. TPA ensure integrity and availability of data and achieves digital forensics and credibility on cloud computing. Third party auditing protocol should satisfy the properties performance, privacy protection i.e. owner's data confidential against auditor, dynamic updates in the cloud and support for batch auditing for multiple owner and multiple clouds. A remote integrity protocol will allow auditor to check data integrity on the remote server. Cloud customers demand the ability to run their own security audits, ensure that proper security measures are always in place and be able to control their security policies inside their own private cloud [21].

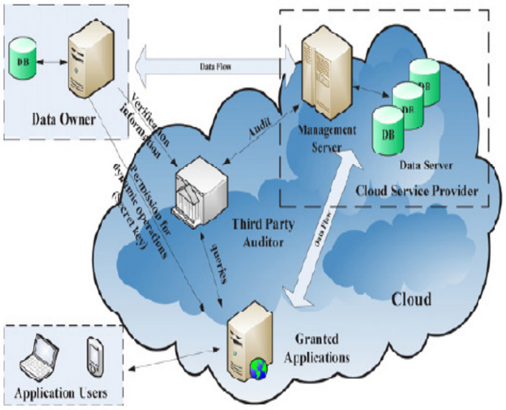

Fig1 Audit System architecture for Cloud computing

vii. *Implication C7:*IaaS cloud providers should ensure that virtualized infrastructure is secure against vulnerabilities. Attack methods such as phishing, fraud, sql injections, cross-site scripting launch attacks on hypervisor. Attacker might successfully escape virtual machine environment. In cross tenant access of virtual machines an attacker may rent a virtual server as a service customer and use administrative rights to analyze configuration, patch level etc to attack other customer's images. An attacker may also use manipulated image to provide back-door access[20]. Service provider should guard against above mentioned vulnerabilities.

viii. *Implication C8*:Data abuse in the cloud can be avoided by encryption of data. Users storing the data in cloud would like to maintain the privacy of their data and are concerned about the security of the data. There is lot of research on protecting data through encryption based techniques. Encryption however increases response time due to computational complexity of data encryption/decryption. Pay per use model of cloud requires lower computational cost and lower computation overhead to the data owner. Data encryption also poses a challenge of finding required tuples to execute query over encrypted data. A filtering mechanism is used [9] that selects required tuples depending upon data content. It results query response time in positive way but hampers privacy performance issue. With homomorphic token with distributed verification of erasure-coded data, a scheme [6] achieves integration of storage correctness insurance and data localization. A progressive elliptic curve encryption scheme which allows a piece of data to be encrypted multiple times using different keys such that the final cipher text can be decrypted in a single run with a single key [17]. A model in [11]involved attribute based encryption and proxy re-encryption, which does not require key redistribution and data re-encryption.

ix. *Implication C9*:CSP can be bound to a specific SLA which assures e.g. data availability or backup procedures to the costumers. External threats are increasing globally, with economic losses from all types of disasters escalating rapidly. Protecting business-critical data, applications and operations against downtime and disruptions is crucial for every organization. A CSP should provide a cloud-independent backups as well as cross-cloud disaster recovery. Application-specific SLA's must be defined in order to identify which applications should be moved to cloud-based DR.Changes to management and legacy recovery processes must be documented.

x. *Implication C10*: Provenance is a meta-data describing the derivation history of data. Provenance is crucial in order to enhance confidentiality, reliability, transparency of digital objects in cloud. It is required to verify the authenticity or identity of data. In scenarios that haveclear requirements for maintaining the provenance of data, including eScience and healthcare, where assurance in the quality and repeatability of results is essential. In addition, clouds have their own application for provenance: the identification of the origins of faults andsecurityviolations[22]. By analyzing and utilizing provenance, it is possible to detect various data leakage threats and alert data administrators and owners; thereby addressing the increasing needs of trust and security for customers' data. Cloud provenance will help in providing trustworthy cloud to support automated management services with incident provenance describing, for example, the root cause of an incident. Secure provenance is vital to the success of data forensics in cloud computing, yet it is still a challenging issue today. At present cloud provenance is provided through log and audit data. A rule-based data provenance tracing algorithm, which trace data provenance to detect actual operations that have been performed on files, especially those under the threat of leaking customers' data isdeveloped in [23].

V CONCLUSION

Ensuring security and privacy for data management and query processing in the cloud is critical for better and broader adoption of the cloud. Storing the data in the cloud relieves the users from the complexities of hardware management. But cloud services are vulnerable to attacks and will bring losses to users because their data is stored on data centers outside enterprise. This paper mainly focused on data security issues in cloud computing and various measures adopted to make the data secure in cloud environment. In order the increase the integrity, availability, privacy protection in cloud computing various schemes are proposed like data as a service, encryption of data, cloud data audit etc. Traditional cryptographic techniques will be of no use in cloud computing because it increases the computational cost. Various new cryptographic techniques like elliptic curve encryption, homomorphic tokens and verifiable tags can be used in cloud computing to increase the data integrity. Some research is also focused on Third party auditing which increases data integrity and data availability. In this paper we have outlined overview of current data security concerns in cloud and discussed various potential implications for data security in cloud computing. In a subsequent work more elaborate survey can be undertaken. We also plan to create a framework in which data security risk assessment can be taken.

REFERENCES

[1] Lisa J. Sotto, Bridget C. Treacy, and Melinda L. McLellan, Privacy and Data Security Risks in Cloud Computing, Electronic commerce and law report, February 3, 2010, http://www.bna.com

[2] Dean, David &Saleh, Tamim: Capturing the Value of Cloud Computing. HowEnterprises Can Chart Their Course to the Next Level. The Boston ConsultingGroup, 2009, p. 1, available at: www.bcg.com/documents/file34246.pdf.

[3] Mell, Peter & Grace, Tim: The NIST Definition of Cloud Computing. NationalInstitute of Standards and Technology, 2009, available at:http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc.

Balboni, Paolo, Mccorry, Kieran & Snead, David: Cloud Computing – Key Legal Issues. In: Cloud Computing Risk Assessment. European Networks and Information Security Agency (ENISA), 2009, p.97111,availableat:http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing riskassessment/at_download/fullReport.

[5] NASSCOM Report, Data Protection challenges in cloud computing, Study Report 2010.

[6] K. ValliMadhavi, Data storage security in cloud computing for Ensuring Effective & Flexible Distrubted System, International Journal of Electronics Communicationa and Computer Engineering, Volume 3, Issue(1) NCRTCST, ISSN 2249-071X

[7] Cong Wang, Qian Wang, and Kui Ren, Ensuring Data Storage Security in Cloud Computing,

[8] D. Agrawal et al., "Database Management as a Service: Challenges and Opportunities," *IEEE 25th Int'l Conf. Data Engineering* (ICDE 09), IEEE CS Press, 2009, pp. 1709–1716.,

[9] H. Hacigumus, B. R. Iyer, C. Li, and S. Mehrotra, "Executing SQL over encrypted data in the database service provider model," in Proc of the ACM SIGMOD Conf., 2002.

[10] Mohammad, et al, Using Multi Shares for Ensuring Privacy in Database-as-a-Service, Proceedings of the 44th Hawaii International Conference on System Sciences – 2011

[11] Yanjiang Yang, Youcheng Zhang, A Generic Scheme for Secure Data Sharing in Cloud, 2011 International Conference on Parallel Processing Workshops.

[12] G. Zhao, C. Rong, J. Li, F. Zhang, and Y. Tang. Trusted data sharing over untrusted cloud storage providers. Proc. 2nd IEEE International Conference on Cloud Computing Technology and Sciene, pp. 96-103, 2011.

[13] Subashini S, Dr. Kavitha V, A Metadata Based Storage Model For Securing Data In Cloud Environment, 2011 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery

[14] Kan Yang, Student Member, IEEE, XiaohuaJia, Senior Member, IEEE, An Efficient and Secure Dynamic Auditing protocol for Data storage in Cloud, IEEE Transactions on Parallel and Distributed Systems

[15] Gartner: Seven cloud-computing security risks. InfoWorld. 2008-07-02 http://www.infoworld.com/d/security-central/gartner-seven-cloudcomputing-security-risks-853.

[16] Yan Zhua, HongxinHuc, Gail-JoonAhnc, Stephen S. Yauc, Efficient audit service outsourcing for data integrity in clouds, The Journal of Systems and Software 85 (2012) 1083– 1095

[17] Gansen Zhao, ChunmingRong, Jin Liz, Feng Zhang and Yong Tang, Trusted Data Sharing over Untrusted Cloud Storage Providers, 2nd IEEE International Conference on Cloud Computing Technology and Science.

[18] Telecommunication Industry Association, TIA-942: Data Center Standards Overview , http://tiaonline.org

[19] Safwan Mahmud Khan and Kevin W. Hamlen,A Data Ownership Privacy Provider Framework in Cloud Computing

[20] Bernd Grobauer, Tobias Walloschek, and Elmar Stocker, " Understanding Cloud Computing Vulnerabilities", Copublished by the IEEE Computer And Reliability Societies 1540-7993/11/ © 2011 IEEE March/April 2011

[21] R. Chow, P. Golle, M. Jakobsson, R. Shi, J. Staddon, R. Masuoka, and J Molina, Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control, in Proceedings of the 2009 ACM Cloud Comuting Security Workshop

[22] Imad M. Abbadi and John Lyle, Challenges for Provenance in Cloud Computing

[23] Olive Qing Zhang, Ryan K L Ko, Markus Kirchberg, Chun HuiSuen, Peter Jagadpramana, Bu Sung Lee, How to Track Your Data: Rule-Based

Data Provenance Tracing Algorithms, 2011 Third IEEE International Conference on Cloud computing Technology & Science