Data Security for Cloud computing Storage System

Jayshree Shinde
PG student
Department of IT Sinhgad College of Engineering
Pune, Maharashtra, India-411041
Jayushinde0009@gmail.com

U. R. Godase
Associate Professor
Department of IT Sinhgad College of Engineering
Pune, Maharashtra, India-411041
urgodase.scoe@sinhgad.edu

***Abstract*— Cloud computing is a forthcoming revolution in information technology industry because of its performance, accessibility, low cost and many other luxuries. It provides storage for data and faster computing to customers over the internet. That is why companies are reluctant to deploy their business in the cloud even cloud computing offers a wide range of luxuries. Security of data in the cloud is one of the major issues which acts as an obstacle in the implementation of cloud computing. Considering to the high storage capacity and also qualified personnel it becomes difficult pertaining the management of such huge amount of data . Some of the facilities like Storage-as-a-Service offered by cloud service providers help organizations to outsource their data to get saved on Remote servers. Thus, such services like Storage as a service helps in reducing the cost of maintenance and also helps to reduce the SaaS burden of large local data storage at the organization's perspective. A storage scheme which is cloud based allows the owner of data to get benefit from the extra facilities given by the cloud service providers and helps in achieving indirect mutual trust in them. There are two important features first is, it allows the owner to outsource sensitive data to a cloud service provider, and it ensures that only authorized users receive the outsourced data. Secondly there is a facility which activates the indirect mutual trust in between data owner and the cloud service provider.**

Keywords- Cloud Computing, Data Security, Storage-as-a-Service, Mutual Trust, Access Control.

I. INTRODUCTION

Cloud is an Internet-based computing technology, where shared resources such as storage, platform, software and information are provided to customers on demand. Cloud computing has received considerable attention from both academia and industry due to a number of important advantages including: cost effectiveness, low overhead management, immediate access to a wide range of applications, mouldability to scale up and down information technology capacity, and mobility where user can access information wherever they are, not present at their desk. Cloud computing is a distributed computational model over a large pool of shared virtualized computing resources e.g. memory, processing power, storage, applications, services, and network bandwidth. In the current era of digital world, many organizations produce a large amount of data including personal data, E-health records, and data related to Finance. The local management of such large amount of data is problematic and costly as a reason of high storage capacity and also requirement of qualified personnel. Therefore, Storage-as-a-Service provide by cloud service providers emerged as a solution to mitigate the burden of large local data storage and reduce the cost of maintenance by means of outsourcings storage of data. Since the data owner physically delivers sensitive data to a remote Cloud Service Provider there are some concerns factors like confidentiality, integrity and access control of the data. The feature such as can be supported by the owner via encrypting the data before outsourcing to remote servers. For checking data integrity over cloud servers, Development and Researcher activity teams have newly prescribed provable data possession technique to validate the intactness of data stored on remote sites. For the access control, the possessor of the informational data encrypted the data under certain key, which is shared only with the legitimate users who have rights to access the data. The unauthorized users, including the CSP, are not access the data because they do not have the decryption key. In this theme, we have brought up a idea that lightens important facts and issues related to outsourcing of the storage of data which includes mutual trust, newness, dynamic data, and access control. The data which is stored on cloud server which is not just accessed by appropriate users also updated by owner. After the updating is over, sanctioned users can receive the updated version of the data, a new method is necessary to identify if received data is stale. The Trust between CSP and data owner can be treated as a secondary issue. A mechanism introduced to determine the dishonest party. This access control is taken into account which helps owner to give permission or deny to the outsourced information in form of data.

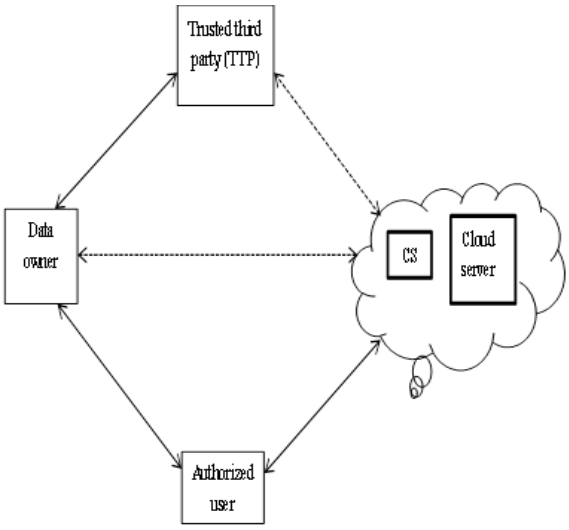

Figure 1: Cloud data storage service

II. RELATED WORK

Yu S, Wang C, Ren K and LouW [1] have proposed a methodology that helps the data owner to achieve fine-grained access control on files stored in Cloud Servers. This method supports user responsibility. It enables the data owner to delegate most of computation intensive tasks to cloud servers without disclosing data contents. The methodology combines three advanced cryptographic techniques such as KP-ABE, PRE and lazy re-encryption. Here any data file is associated with a set of attributes, and assign each user demonstrative access structure which is defined over these attributes. A PRE scheme allows the proxy, given the proxy re-encryption key, to translate cipher texts under one public key into cipher texts under another public key and vice versa. Main advantages in the method are that the data owner can store data files and also run his own code on Cloud Servers to manage his private data and achieve security goals like support and basic operations like user authenticate/revoke. The only drawback is that the Cloud Servers will enable trace to search as much confidential information as possible based on their valuable inputs

Goyal V, Pandey O, Sahai A,Waters [2] have proposed a new cryptosystem for fine-grained sharing of encrypted data that called Key-Policy Attribute-Based Encryption (KP-ABE). This scheme uses a set of attributes to describe the encrypted data and builds a access policy in user's private key. If attributes of the encrypted data can satisfy the access structure in user's private key, then the user can obtain the message through decrypt algorithm. In a policy of key and attribute-based encryption (KP-ABE) system, cipher texts are labeled by the sender with a set of relevant informational attributes, while user's private key is generated and issued by the trusted attribute authority which captures an policy that is also regarded as access structure, that specifies the type of cipher texts the key can specify and decrypt. KP-ABE schemes are made in accordance with structured organizations that are inline and ready to get compared with particular documents.
This method helps to share stored data in the fine grained level. The drawback in KP-ABE system is that encryptions and conversion to cipher ext is not allowed to generate the access policies and its relevant dependency on the well-defined access control mechanism. On each user revocation, a new access policy is defined, which lead to a situation where it is not possible to construct further more access structure.

Goh E-J, Shacham H, Modadugu N, Boneh [7] have proposed design of security mechanism that improves and guarantees the security and trust of a networked file system without making any alterations to the file system. SiRiUS is layered over existing file systems such as NFS to provides security. To achieve the access control, SiRiUS attaches each file with a meta data file which contains that file's access control list (ACL), containing each entry of which is the encryption of the file's file encryption key (FEK) using the public key of an authorized user. SiRiUS lets it behave as a local file system with the preset standard hierarchical view of directories and files. SiRiUS is defends against version rollback attacks and it decrease traffic generated by network by providing a facility of random access within files. Rea-write and also read only access is supported by it. Disadvantages are
1. Scalability: collaboration among bigger groups SiRiUS proves to be inefficient. Whenever, user is revoked, FEK and FSK are updated and new keys are encrypted with public key of the individual user, making it for dynamic user sets. Writes and reads to any location in a file must take amount of time that is comparable.
2. Key Management. The rapid increase of keys used by various applications creates a so called usability nightmare.

Benaloh J, Chase M, Horvitz E, Lauter K [5] have proposed an efficient system that allows patients both to share partial access rights with others, and to perform searches over their records. A centralized storage system with social order was developed for sharing the PHR. It is not possible to completely believe a certificate authority (CA) for managing the storage which guides to key security breach. To overcome this breach, users in the system are led into a classification as public and personal domains. Personal domain manages the information that is personal to the patient. This information is accessible by the owner of the data. Public domain stores and comprises of various types of informational data. An authority is assigned to each type of information. PHR is used by ABE.[5] Thus, personal domain is managed and in hand of the Data Owner that is comprised of KP-ABE and the public domain is managed by multiple attribute authorities which use Multiple Authority - Attribute Based Encryption (MAABE). The attribute authority is only the entity that is responsible for revoking and granting permission to the users.To update the access policies the attributes present in the cipher-text are changed and modified.This method is helpful because the patient can easily grant access to a category, without knowing all the types included in it. Doctors can add subcategories with arbitrary names. Limitation of its hierarchy is like there is only one way to partition the record.

Sultan Ullah, Zheng Xuefeng and Zhou Feng [3] have proposed framework of access control for cloud computing is introduced in this paper, which provides a multi - step and multifactor authentication of a user. The model proposed is well-organized and provably secure solution of access control for externally hosted applications. The validation of the user is a multistep process, and after the successful validation of the user will only access the data file store by the owner, in a confidential means by the implementation of the digital certificate. Advantages is that it Increase the confidentiality and integrity of the data using biometric data disadvantages are there associated with Key management Complexity and difficulty in revocation.

Kevin Fu. U. Mass. Amherst. Seny Kamara [4] have proposed an efficient key distribution for Secure Distributed Storage and solves the key rotation problem. Here member states are given to the authorized users. Using key derivation function encryption key is generated for that particular member states. Encryption keys are separated from the member state so that the key is pseudorandom for any particular member state. Lazy revocation method is used here which postpones the re-encryptions till the next write access is performed so that the extra re-encryptions are eliminated. This method significantly reduces the bandwidth requirements of a content publisher and reduces lack of pseudo randomness in key rotation. The main disadvantage is that after certain number of revocations Key derivation function will not be able to generate the new member states, thus reducing its practicability to a limited number of revocations.

Geron E, Wool A [6] have proposed a cryptographic remote storage system, which avoids the use of public key encryption for the purpose of speed and efficiency in cryptographic operations. Here for each new user, trusted agent generates an encryption key by using system master key stored locally on it. CRUST [6] maintain file in blocks each encrypted with a separate key. The major advantage of this method is that, only the updated block of a file is re-encrypted to avoid unnecessary cryptographic operation. But it uses trusted agent, thus making it highly dependent on the trusted third party. And it suffers from key management problem that is Separate encryption key for each block of a file increases key management burden on each user as well as on the trusted agent.

Saman Zarandioon1, Danfeng (Daphne) Yao2, and Vinod Ganapathy [2] have proposed a user-centric privacy preserving cryptographic access control protocol called K2C (Key To Cloud) that enables end-users to securely store, share, and manage their sensitive data in an unstrusted cloud storage anonymously.K2C is realized through our new cryptographic key-updating scheme, referred to as AB-HKU. K2C is scalable and supports the lazy revocation, and it avoids trusted third party. But it requires guaranteed availability of the data owner until all of the legitimate users update their keys

Zeeshan Pervez, Asad, Masood Khattak [11] have proposed method helps to achieve access control technique that is fine-grained to the outsourced contents of data present in the cloud. Key distribution and management process is done here without analyzing any confidential information about the secure data contents, so that it provides more privacy to the data contents stored. User revocation is achieved by changing one attribute associated with the key, here there is no need to modify the entire access control policy and this enables authorized users to update their decryption keys during each user revocation. These are made possible by combining cipher text policy attribute based encryption and key distribution methods. In this case the main advantage is that the owner should not remain always online to distribute new decryption key among the legitimate user and the legitimate users are allowed to update their secret key after each user revocation without interacting with the owner. Cloud server should not be able to learn any information about the contents of the outsourced data. Desired symmetric encryption methods can be used along in this method. But there is need for a secure channel to transfer the private key between the owner and the user.

III. PROPOSED SYSTEM

ALL THE FEATURES THAT ARE PART OF IMPLEMENTATION AND DESIGN OF A SCHEME THAT IS CLOUD BASED ARE LISTED AS BELOW:

- It gives a facility so that owner of data can outsource the data to CSP, and perform full dynamic operation, i.e., it supports operations such as modification, insertion, deletion, and append
- It helps in keeping us updated and which is also called to be maintaining newness property i.e we get the latest version of data that is outsourced.
- It also establishes trust that is indirect mutual trust in between the CSP and data owner because each party resides in a different trust domain; and
- It enforces the access control for the outsource data.

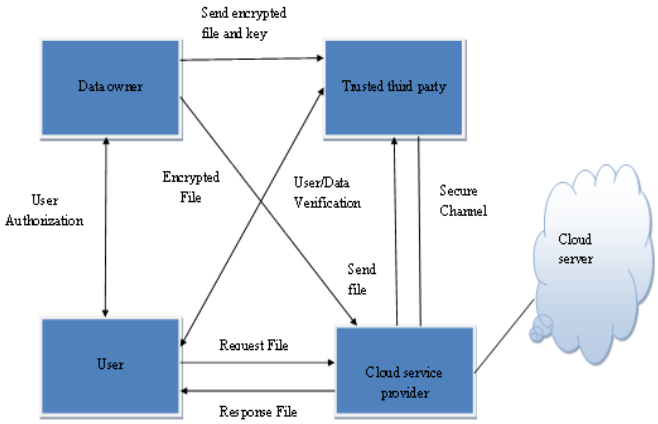

Figure 2: System Architecture

A. *System Componant Relation*

This model of cloud computing storage considered in this scheme consists of four major components as follows.

1. A data owner that can be associated with organization which generates data that is sensitive to be stored in the cloud and made readily available for user;

2. A CSP who takes a responsibility of cloud servers and provides facilities that are charged on its infrastructure to store the owner's data and files and are available for users that are authorized;

3. Authorized user-a set of owner's clients who have the possible way to get access to the remote data for some or the other purpose; and

4. A trusted third party (TTP), is an entity who is known to all and trusted by other components of the system, and has capabilities to detect/specify dishonest parties. The authorized users and data owners have mutual trust relation with the CSP. Thus, the TTP is only used to grant and enable indirect trust which is mutual in between these three components. There is a direct trust relation relationship between the authorized users and data owners.

B. Outsourcing, Updating, Accessing Data

To achieve confidentiality, the owner has a provision made in such a way that it encrypts the data before sending to cloud based servers. After the data is being outsourced, the owner has a provision so that it can interact with the Cloud service provider to perform operations that are dynamic on the saved file. To get access to this data, the legitimate user can sends a access request of data to the Cloud Service Provider, and then it gets back the data file which is in an encrypted form that can be decrypted using a secret key generated by the authorized user. Most of the cases it is taken into consideration that the interaction between the authorized users and the owner to identify and verify their authorized identities is completed to a greater extent. Further the Trusted Third Party does not depends on any other entity, it is a independent entity, Also , any possible corruption of information towards the TTP must be prevented to keep the outsourced data private. The Trusted third Party and the Cloud Service Provider are always online while the data possessor is not always in online mode .The users that are authorized are able to access the information from the Cloud Service Provider and even when the data possessor is at offline state.

C. Threat Model

The Cloud service Provider is not trusted and thus the and integrity and confidentiality of informational-data stored in the cloud may be at a severe risk. In order to maintain a status and economic reasons the Cloud Service Provider may hide loss of data, or recover the storage by revoking the data that has not been or is rarely used., The CSP may totally avoid in order to save the computational resources and the data-update requests, or take into consideration and execute only few of it or may not. This is why, the Cloud Service Provider may return stale or damaged data for any access request from the users that are authorized. Moreover, the Cloud Service Provider may not give the access rights created by the data possessor, and permit access that is unauthorized for the misuse of secret data. On the Contrary, the users that are authorized and the data possessor may collude and falsely accuse the Cloud Service Provider for getting a certain amount of reimbursement. They may also dishonestly claim that integrity of the data over servers of the cloud has been violated or breached, or the Cloud Service Provider has returned a stale file that does not match the most recently alterations and modifications that are prescribed by the owner of data.

D. Security Analysis

• Confidentiality: The outsource data are kept secret; the data owner created an encrypted version of the data file $\breve{f}$ .The encryption of a file is done using a secret key K generated by owner, where K is accessed only by the data owner and authorized users. TTP send the encrypted file to CSP and CSP provide the security to the cloud storage system.

• Data integrity violation: A small part of the owner's work is delegated to the TTP to reduce the storage overhead and lower the overall system computation. For the TTP to resolve disputes that may arise regarding data integrity it computes and locally stores hash value for the encrypted file $\breve{f}Htt$. The TTP sends encrypted file $\breve{f}$ to the CSP. The TTP keeps only $\breve{f}Htt$ and ENCs (K) on its local storage.

• Access control: The owner creates ENCs (K) and only authorized users can decrypt ENCs (K) using S and get key K to read the outsourced data, and thus the access control is achieved in the proposed scheme. The unauthorized user can't decrypt the encrypted file because they don't have the secret key for the decryption of the file F.

• Dishonest owner/user: If the owner/user falsely accuses the CSP regarding data integrity, the TTP performs cheating detection procedure. In this procedure, TTP retrieves encrypted file from CSP and computes the temporary hash value $\breve{f}Htt$ and compares $\breve{f}Htt$ and $\breve{f}$Htemp. If compare $\breve{f}Htt$ =!$\breve{f}$Htemp, then file $\breve{f}$ has not been corrupted on the server and owner/ user is dishonest.

• Dishonest CSP: During the data access phase of the proposed scheme, the authorized user receives the encrypted file $\breve{f}$ from the CSP and $\breve{f}Htt$ from the TTP. The authorized user computes hash of encrypted file $\breve{f}$Hu and compares and $\breve{f}$Hu. If F Http =! FHu, a report is issued to TTP to determine the dishonest party. The TTP retrieves encrypted file from CSP and computes the temporary hash value $\breve{f}$Htemp and compares $\breve{f}Htt$ and $\breve{f}$Htemp. If $\breve{f}Htt$ =!$\breve{f}$Htemp then has been corrupted on the server and CSP is dishonest.

IV. CONCLUSION AND FUTURE SCOPE

Cloud storage security and privacy risks are identified and a Security as a Service design is proposed which could securely access data from CSP. The motivation behind this research lies in the fact that for many organizations the final barrier to adopting Cloud computing is whether it insufficiently secure.

After analyzing cloud storage security and privacy risks, data protection requirements and security applied to current cloud storage services (Amazon's Cloud Drive and

Drop Box), the security for cloud storage services is proposed. Thus proposed design of the service meets most of the defined security and privacy requirements

ACKNOWLEDGMENT

I hereby take this opportunity to express my heartfelt gratitude towards the people whose help was very useful for the completion of my research work on the topic of " Data Security for Cloud computing Storage System " It is my privilege to express sincerest regards to the project Guide Mrs. U. R. Godase, for her valuable inputs, able guidance, encouragement, whole-hearted cooperation and constructive criticism throughout the duration of my project work. I deeply express my sincere thanks to our Co-guide Mr. A. N. Bhute for his encouragement and support.

REFERENCES

[1] Zeeshan Pervez, Asad, Masood Khattak, Sungyoung, Young-Koo Lee, A M Khattak, S Y Lee, Y K Lee SAPDS: self-healing attribute-based privacy aware data sharing 2012. Yu S,Wang C, Ren K, LouW(2010) , Issue 1, pp 431-460

[2] Saman Zarandioon1, Danfeng (Daphne) Yao2, and Vinod Ganapathy K2C: Cryptographic Cloud Storage with Lazy Revocation Anonymous Access. 96, 2012, pp 59-76

[3] Sultan Ullah, Zheng Xuefeng and Zhou Feng TCLOUD: A Multi – Factor Access Control Framework for Cloud Computing. International Journal of Security & Its Applications; Mar2013, Vol. 7 Issue 2, p15

[4] Kevin Fu. U. Mass. Amherst. Seny Kamara. Johns Hopkins University.Key Regression: Enabling Efficient Key Distribution for Secure Distribute Storage.2005

[5] Benaloh J, Chase M, and Horvitz E, Lauter K (2009) Patient controlled encryption: ensuring privacy of electronic medical records. In: Proceedings of the 2009 ACM workshop on cloud computing security, CCSW ’09. ACM, New York, pp 103–114

[6] Geron E, Wool A (2009) Crust: cryptographic remote untrusted storage without public keys. Int J Inf Secur 8:357– 377

[7] Goh E-J, Shacham H, Modadugu N, Boneh D (2006) Sirius: Securing remote untrusted storage. In: Proceedings of the fifth workshop on the economics of information security (WEIS 2006)

[8] Goyal V, Pandey O, Sahai A,Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on computer and communications security, CCS ’06. ACM, New York, pp 89–98

[9] Amazon.com, “Amazon Web Services (AWS),” Online at http://aws.amazon.com, 2008.

[10] J.Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J.Quisquater, “Efficient remote data possession checking in critical information infrastructures,” IEEE Trans. on Knowl. And Data Eng., vol. 20, no. 8,2008.

[11]John Bethencourt Carnegie Mellon University“Ciphertext-Policy Attribute-Based Encryption” Security and Privacy, 2007. SP '07. IEEE Symposium 20-23 May 2007 321 - 334

[12] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, “Scalable and efficient provable data possession,” in Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks, 2008, pp. 1–10.

[13] C. Erway, A. K¨upc¸ ¨ u, C. Papamanthou, and R. Tamassia, “Dynamic provable data possession,” in Proceedings of the 16th ACM Conference on Computer and Communications Security, 2009, pp. 213–222.

[14] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, “Enabling public verifiability and data dynamics for storage security in cloud computing,” in Proceedings of the 14th European Conference on Research in Computer Security, 2009, pp. 355–370.

[15] A. Juels and B. S. Kaliski, “PORs: Proofs of Retrievability for large files,” in CCS’07: Proceedings of the 14th ACM conference on Computer and communications security. ACM, 2007, pp. 584–597.

[16] H. Shacham and B. Waters, “Compact proofs of retrievability,” Cryptology Print Archive, Report 2008/073, 2008, http:// eprint.iacr.org/.