Data Security And Its Techniques In Cloud Storage – A Review

Priyanka V. Mogre
ABHA Gaikwad-Patil College of Engineering,Nagpur

Girish Agarwal
ABHA Gaikwad-Patil College of Engineering,Nagpur

Pragati Patil
ABHA Gaikwad-Patil College of Engineering,Nagpur

Abstract

The purpose of this review paper is to provide information regarding techniques that could be beneficial for the platform cloud computing by providing security in cloud storage. The main issue of cloud storage is that data must be secured, by using various security tools; also the privacy of the data must be maintained. Various researchers researched in the field of security and privacy for data stored in cloud. By referring various researchers work done in the field of data privacy and security, in this paper we will study, analyse and review techniques for data security in cloud storage.

Keywords: Cloud storage, Cloud computing, Cloud data storage security, trusted third party (TTR).

1. Introduction

The architecture of many IT companies is built up by cloud computing. As large number of people working in enterprise or any such kind of technical field comes across various data which as to be stored for their further use. To store small amount of data within the computer is possible but if the data stored is of large size, it becomes difficult to store such large data within the computer. For providing storage of such large data we require cloud server were the data can be stored as much as user want.

Cloud Computing precisely is defined as a service which is delivered over internet (Network) using various computing resources. Most IT companies are forced to spend a significant portion of their time on frustrating implementation, maintenance, and upgrade a project that too often does not add significant value to the company's bottom line. Increasingly, IT teams are turning to cloud computing technology to minimize the time spent on lower-value activities and allow IT to focus on strategic activities with greater impact on business. A cloud can either be a private cloud or a public cloud. Private cloud is the one where there is a data centre that supplies services to the limited number of people. Whereas, public clouds are that which supplies services to anyone on the internet. A virtual private cloud is formed when the service provider uses public cloud resources to create private cloud.

The characteristics of cloud computing are,

i. It is *virtual*, which means there are large numbers of servers which are placed along data centre. This server becomes massive pool of resources this pool is divided into various multiple virtual servers which lead to the creation of 'cloud'.
ii. It is *flexible and scalable*, which means it gives whatever the user need within a moment. It also spins up the server in a moment and take it down just as easily.
iii. It is *open or closed*, in open cloud it can be easily moved around without been locked into one provider or a closed, proprietary technology.
iv. It can be *secured*, for maintaining the security there cloud be a creation of private cloud on the working hardware, but an appropriate security measures must be put on these cloud.
v. It can be *affordable*, a good cost saving could be made on public cloud, whereas in virtual servers runs on physical servers that are shared with other customers.

The limelight that cloud computing always requires in any IT companies which makes them think what IT always need is a way to increase capacity or add capabilities without investing on any new infrastructure, or licensing new software, or training any new personnel. Cloud computing encompasses pay-per-use service that in real time over the internet.

As from above introduction it is clear that cloud computing is all about storing more of the data or material somewhere in the cloud and using whenever required by the user and maintaining PC's or servers with less amount of data which make it easy to access whenever user or client is in a search of data created or stored. The main benefit of cloud computing is being to access data anywhere within an internet connection.

Cloud storage is a networked online storage where data is stored in a virtualized pool which is generally hosted by the trusted third party (TTR), TTR operates large data centres. The operation of data centre is played at the background; it virtualizes the resources according to the requirements of the customer and exposes them as a storage pool, which the customer can themselves use to store files or data objects. As per the characteristics of cloud storage with respect to cloud computing, it does not have same characteristics as that of cloud computing regarding in terms of agility, scalability, elasticity and multi-tenancy. Cloud storage is made up of many distributed resources with the potential for the economics of scale. Rather than cost, its benefits are outsourced operations, simple, enterprise feature for smaller users like high availability, security, data protection, privacy, etc. The interface to the cloud storage will be visible to the user and storage manager. This is a sort of software application that runs locally and sends data to the cloud. It look like a network mounted drive which could be integrated into a back up or archive application. As in the way the storage computer is connected to one another there is no limit in storing the large amount of data on cloud computing environment, because as soon as one web server fills up it span across automatically to the next server that is connected in the series so essentially data will be automatically distributed across multiple web server.

2. Structure and strategy of literature review:

The sources for literature review are not only limited on references and papers published but also referred online material by using various search engines. For making this review paper we referred various E-books on security and cloud computing along with various survey papers related to my research.

3. DISCUSSION

In the following section will be discussing about the terminologies, technology, issues of security and privacy of cloud computing and cloud storage infrastructure.

As per the discussion point of view cloud computing and cloud storage as been discussed in the above introductory part. The following terminologies discussed below will help to understand the concept more clearly.

1. *Data storage security*: storing or recording of the information on the computer or eventually on similar devices is called Data storage (DS). It could also be said as that the information's or data the computer "knows" or is able to know.DS is trying to combine with storage security to provide more robustness in cloud data storage and forms cloud data storage security(CDSS). As many users are unaware of how the data are stored in the cloud because, the recorded information is not visualized by any user hence the risk of security rises. To decrease this risk only service can be provided is cloud data storage service provider (CDSSP).

2. *Cloud service provider*: The cloud storage provider is a third party company that offers end-users to save data to an off-site storage system. Instead of storing the data to local hard drive or any local storing device the data is stored on a remote data centre. Any authorized person can access these data from anywhere having internet connection.

3. *Provable data possession (PDP)*: Provable data possession is a technique to verify the outsourced data with least computation, communication, and storage overhead. Ateniese et al. have formalized a PDP model. In that model, the data owner Pre-processes the data file to generate some metadata that will be used later for verification purposes through a challenge response protocol with the remote/cloud server. The file is then sent to be stored on an untrusted server and the owner may delete the local copy of the file. Later, the server demonstrates that the data file has not been deleted or tampered with by responding original data owner or other trusted entity that shares some information with the owner[17].

Various researchers undergo different schemes for PDP. The above PDP scheme had one drawback that it only support to PDP schemes that focus on a *single copy* of the file and provide no proof that the CSP stores multiple copies of the owner's file[17].

After the drawbacks of the schemes author worked on other schemes and proposed two schemes for PDP. Two Efficient Multi-Copy Provable Data Possession (EMC-PDP) protocols, and prove the security (*correctness* and *soundness*) of our protocols against colluding servers. Extensive performance analysis — which is validated through implementation and experimental results — illustrates the efficiency of our protocols. Curtmola et al. [15] proposed a Multiple-Replica PDP (MR-PDP) scheme, which is the only attempt in the literature that creates multiple replicas of owner's file and audit them [17].

4. *Cooperative provable data possession (CPDP):* In order to prove the integrity of data stored in a multi-cloud environment, we define a framework for CPDP based on interactive proof system (IPS) and multi prover zero-knowledge proof system (MPZKPS). Homomorphic verifiable response is the key technique of CPDP because it not only reduces the communication bandwidth, but also conceals the location of outsourced data in the distributed cloud storage environment [2].

5. *Security and Privacy*: The best part for maintaining security is cryptography. Cryptography provides various secure techniques in presence of third party for communication purpose. It analysis the protocol that overcomes the influences of adversaries which are related to various aspect in securing information such as data confidentiality, data integrity, and authentication. In this technique the information in converted with the help of Encryption technique in an unreadable format so that only the authorized person can access these information with the help of private key provided by encryption technique. With the help of these private key the encrypted data is converted to the readable form of data this technique is called decryption, which makes the data in the readable format.

   It is beneficial if these techniques are used in cloud storages services where the data are stored from various resources. To maintain the security of these data cryptography plays major roles. The other way of using cryptography is, it is easy to use and understand.

6. *Service Security:* In these scenarios most of the security threats are possible at the point of service provision and this could include the actual device security at the cloud provider and the storage security used by the provider. Though due to the business nature of the service providers they would be able to provide robust security with the use of state of the art IDS, firewalls and malware protection. Moreover the use of virtualization technology further helps the providers in securing each of the individual users from each other [1].

4. Architecture of cloud storage:

The architecture presented in fig 1. Illustrate how data is stored by the application users or the client in the cloud.

Application users or clients stores the data in the cloud and depends on the cloud for computation of data. The client could either be an individual or an organization or an enterprise.

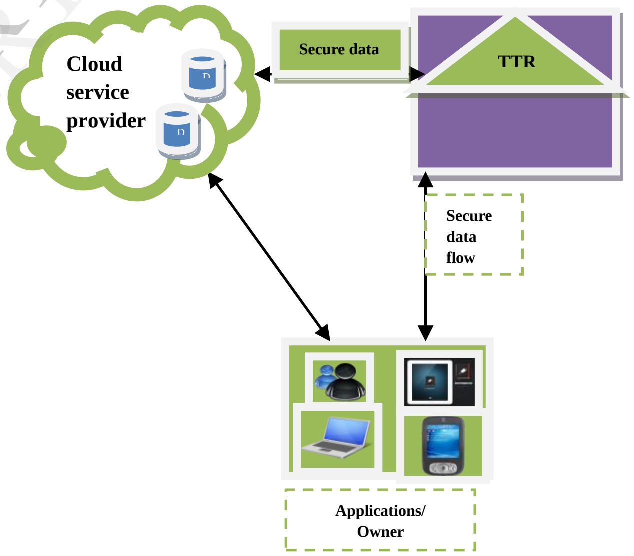

Fig.1: Architecture of Cloud Storage.

Application can be any technological devices which are connected via internet and stores data in remote location. By using cloud service provider (CSP) the owners or the users stores the data in cloud where there are sets of cloud server. Here, to maintain data redundancy a technique is been used called as erasure correcting code which could be further extended to tolerate fault or server crash these is due to growth in user data and the importance. To retrieve data hereafter the client or the user interact or communicate with cloud server via cloud service provider.

As user/client are storing there data/information in a virtual mode i.e. a cloud server and not on the local disk which is visible to the user. As data is stored in the cloud and is not visible to the user the risk of security increases and to maintain the integrity of data becomes difficult.

In such situation user can rely on an optional TTP of their respective choice and can access the data stored in the cloud.

Trusted third party (TTP) is an entity which creates an interaction between two parties. As TTP creates secure zone between two parties it plays its best role in cryptography where with the help of secure key information between only two authorized parties can be maintain and the other person apart from these two people who want to communicate with each other should not be able to know about the information or the data. Here, only the TTP who knows both people can communicate with each other. Hence, it establishes the security to the user and its data stored in the cloud. Here both user and the cloud can trust the TTR model for accessing there data.

5. Conclusion:

In this paper, we focused on how the data is stored in cloud. Also, understood what cloud computing is. What exactly does data stored in cloud means?

Along with the introduction to cloud computing and cloud storage, also referred to the characteristics of cloud computing. Here-by also presented how the structure of our literature survey was built-in. We also discussed various terminologies used by cloud storage along with the technologies including PDP, CPDP for maintaining the data more secure and available. The architecture of cloud storage provided clarity how via CSP information is stored in cloud server and with the use of TTP the interaction between the user and cloud is been secured.

References

[1] Anup Mathew " Survey Paper on Security & Privacy Issues in Cloud Storage Systems", EECE 571B, TERM SURVEY PAPER, APRIL 2012.

[2] Yan Zhu, Hongxin Hu, Gail-Joon Ahn, *Senior Member, IEEE*, Mengyang Yu "Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS.

[3] G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner, Z. N. J. Peterson, and D. X. Song, "Provable data possession at untrusted stores," in *ACM Conference on Computer and Communications Security*, P. Ning, S. D. C. di Vimercati, and P. F. Syverson, Eds. ACM, 2007, pp. 598–609.

[4] Amir Mohamed Talib "Security Framework of Cloud Data Storage Based on Multi Agent System Architecture: Semantic Literature Review", www.ccsenet.org/cis Vol. 3, No. 4; November 2010.

[5] Y. Dodis, S. Vadhan, and D. Wichs, "Proofs of retrievability via hardness amplification," in *TCC'09: Proceedings of the 6th Theory of Cryptography Conference on Theory of Cryptography*, Berlin,Heidelberg, 2009, pp. 109–127.

[6] K. D. Bowers, A. Juels, and A. Oprea, "Proofs of retrievability: theory and implementation," in *CCSW '09: Proceedings of the 2009 ACM Workshop on Cloud Computing Security*, New York, NY,USA, 2009, pp. 43–54.

[7] J. Li, M. Krohn, D. Mazi`eres, and D. Shasha. Secure untrusted data repository (SUNDR). OSDI, 2004.

[8] B. Sotomayor, R. S. Montero, I. M. Llorente, and I. T. Foster,"Virtual infrastructure management in private and hybrid clouds," *IEEE Internet Computing*, vol. 13, no. 5, pp. 14–22,2009.

[9] Blaze, M. A cryptographic file system for unix. In ACM CCS (1993).

[10] Bindel, D., Chew, M., And Wells, C. Extended cryptographic filesystem. In Unpublished manuscript (1999).

[11] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. of SecureComm '08, pp. 1–10, 2008.

[12] Li, J., Krohn, M., Maziered, D., AND Shasha, D. Sundr: Secure untrusted data repository. In OSDI (2004).

[13] J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasurecoded Data," *Proc. 26th ACM Symposium on Principles of Distributed Computing*, pp. 139–146, 2007.

[14] Farkas, C. Huhns, M.N. (2002). "Making agents secure on the semantic web", IEEE Internet Computing (6) (2002)76–79.

[15] Filsinger, J and Lubbes, H. O. (1996). "System security approach for the High Level Architecture (HLA)". In Proceedings of the 14th Workshop on Standards for Interoperability of Distributed Simulation (winter).

[16] R. O. Weichao Wang, Zhiwei Li and B. Bhargava, "Secure and efficient access to outsourced data," in *CCSW '09: Proceedings of the 2009 ACM Workshop on Cloud Computing Security*, New York,NY, USA, 2009, pp. 55–66.

[17] Ayad F.Barsoum and M.Anwar Hasan, "Provable Possession and Replication of Data over Cloud Servers"