IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-4, Issue-3 E-ISSN: 2347-2693

Privacy Preservation and Auditing in Public Cloud: A Review

Nitesh Kumar Namdeo[1*] and Sachin Choudhari[2]

[1*,2] *Department of Computer Science and Engineering, RGPV, India*

www.ijcseonline.org

Received: Feb/22/2016 Revised: Mar/06/2016 Accepted: Mar/14/2016 Published: Mar/31/2016

Abstract- Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing is a general term for the delivery of hosted services over the internet. Cloud computing enables companies to consume the resources and compute their utility rather than building and maintaining computing infrastructure. A cloud database is a database that has been optimized or built for a virtualized computing environment. Since these data-centers may be located in any part of the world beyond the reach and control of users, there are multifarious security and privacy challenges that need to be understood and addressed. Cloud has been prone to various security issues like storage, computation and attacks like Denial of service, Distributed Denial of Service, Eavesdropping, insecure authentication or logging etc. Privacy preservation is main security issue in public cloud. This paper focuses on various security mechanisms that are provided in the enterprises and also discusses few of the common security mechanisms like auditing, authentication, authorization, encryption and access control.

Keywords- Cloud database, Security, Privacy Preservation, Auditing, Authentication, DaaS.

I. INTRODUCTION

An important aim of cloud computing [1] is to provide on-demand access to computational resources on pay-as you-go basis similar to the way in which we obtain services from public utility services such as water, electricity, gas and telephony. Essentially, there are two main stakeholders in the Cloud Computing environments, which are the Cloud providers (service producers) and Cloud customers (service consumers or clients). Cloud customers can be either software/application service providers who have their own service consumers or end users (e.g., organization or businesses) who use cloud computing services directly. A cloud provider is a company or vendor that offers economically efficient cloud services using the hardware and software.

Massive growth in digital data[2], changing data storage requirements, better broadband facilities and Cloud computing led to the emergence of cloud databases .Cloud Storage, Data as a service (DaaS) [3] and Database as a service (DBaaS) are the different terms used for data management in the Cloud. They differ on the basis of how data is stored and managed. Cloud storage is virtual storage that enables users to store documents and objects. Drop box, iCloud [3] etc. are popular cloud storage services. DaaS allows user to store data at a remote disk available through Internet. Cloud storage cannot work without basic data management services. So, these two terms are used interchangeably. DBaaS is one step ahead. It offers complete database functionality and allows users to access and store their database at remote disks anytime from any place through Internet. Amazon's SimpleDB, Amazon RDS, Google's BigTable, Yahoo's Sherpa and Microsoft's SQL Azure Database are the commonly used databases in the Cloud. The data should be kept secured and should not be exposed to anyone at any cost. Confidentiality of data is another security issue associated with cloud computing. The different security issues in cloud are scalability, heterogeneity, Data Intrusion [4], Data Integrity, Non-Repudiation, Confidentiality, access control, authentication and authorization. Based on many discussions with customers and surveys, the following security and integration issues seem to be on many customers' minds. How the cloud will keep data secure and available? How to comply with current and future security and risk management compliance? What types of security services are available through the cloud? How to perform internal and external audits of cloud security? How to automate network, compute, and storage provisioning?

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources(e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. A style of computing where massively scalable IT-related capabilities are provided 'as a service' using Internet technologies to multiple external customers. At its simplest, cloud computing is the dynamic delivery of information technology resources and capabilities as a service over the Internet. How to do on-demand provisioning in near real time from a customer portal to all the infrastructure devices

How to orchestrate among many new cloud tools and existing legacy tools Although most of the surveys show that most customers are concerned about security and integration, most of the successful organizations are taking calculated risks and implementing the cloud with appropriate security measures. As many of you know, nothing can be 100 percent secure, but by knowing the current state, one can apply appropriate security measures to mitigate the risk and grow the business.

Paper is organized as follows. Section II represents background related to cloud computing, privacy and security of cloud database. Section III provides literature survey of privacy preservation and security. Section IV concludes the paper.

II. BACKGROUND

A. Cloud Computing Services

There are four types of cloud computing service models [5]- SaaS, PaaS, IaaS, DaaS.

Software as a service (SaaS)-Saas can be defined as the software that is deployed over the internet. A complete software is available over the cloud any customer can use that software on "pay-as-you-go" basis.[5] The Saas provides on-demand access of software to the clients. One more characteristic of Saas is that it delivers the software in "one to many" model. In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients.

Platform as a service (PaaS)[6]-In platform as a service model, service provider provides hardware and software to the customer which is needed by him to database and web server. Paas is a form help enterprise developers quickly develop software. In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming language execution environment and of cloud computing that holds web- potential write and test customer or employee facing application.

Infrastructure as a service (IaaS)-It is the most basic cloud service model. It provides computers physical or virtual machines and other resources. IaaS clouds often offer additional resources such as a virtual-machine. DISK IMAGE library, raw block storage, and file or object storage, firewalls, load balancers, IP addresses, virtual local area networks and software bundles [7].

Database As A Service (DAAS)-Cloud database is designed for virtualized computer environment. It is not as simple as taking relational database and deploying it over a cloud server.[4] Cloud database as a service has to fulfill all the characteristics of relational database as well as cloud database. There are two terms used for data storage in cloud DaaS(Data as a service) &DbaaS (Database as a service).In data as a service only a space is provided over the cloud to store the data but in database as a service client can store data as well as he can run queries over the data to alter them and get some useful information from the database. Clod database is created over the service provider site. So security should be very high in the cloud database because client has to protect his data from the outsider as well as he has to protect the data from the service provider also. It might be possible that database has some harm from the cloud database provider.

Irrespective of the above mentioned service models, cloud services can be deployed in four ways depending upon the customers' requirements.

Public Cloud: A cloud infrastructure is provided tomany customers and is managed by a third party [8]. Multiple enterprises can work on the infrastructure provided, at the same time. Users can dynamically provision resources through the internet from an off-site service provider. Wastage of resources is checked as the users pay for whatever they use.

Private Cloud: Cloud infrastructure, made available only to a specific customer and managed either by the organization itself or third party service provider. This uses the concept of virtualization of machines, and is a proprietary network.
Community cloud: Infrastructure shared by several organizations for a shared cause and may be managed by them or a third party service provider.

Hybrid Cloud: A composition of two or more cloud deployment models, linked in a way that data transfer takes place between them without affecting each other.

B. Security issues in cloud database:

Data Intrusion: Data Intrusion [6] is another security risk that may occur with a cloud provider. Undesirable alteration of user data may commence due to intrusion. If any intruder can gain access to the account password, then he/she will be able to do any kind of unwanted changes to the account's private documents.

Data Integrity: The stored data in the cloud storage may suffer from enormous damage occurring during the transition operations from or to the cloud storage provider. It is very essential to maintain the integrity of data. The risk of attacks from both inside and outside the cloud provider exists and should be considered.

Non-Repudiation: It guarantees the transmission of message between parties and gives the assurance that someone cannot

deny something.. It ensures that a party cannot deny the genuineness of their signature on a document or the sending of a message that they originated. Non-repudiation is a major concern for data security. Non-repudiation is often used for signatures, digital contracts, and email messages

Confidentiality: The data should be kept secured and should not be exposed to anyone at any cost. Confidentiality [6] of data is another security issue associated with cloud computing.. The users do not want their confidential data to be disclosed to any service provider. But it is not always possible to encrypt the data before storing it in cloud.

Access control
Access management [7] is one of the toughest issues facing cloud computing security. One of the fundamental differences between traditional computing and cloud computing is the distributed nature of cloud computing. Within cloud computing, access management must therefore be considered from a federated sense, where an identity and access management solution is utilized across multiple cloud services and potentially multiple CSPs. Access control can be separated into the following functions:

Authentication: An organization can utilize cloud services across multiple CSPs, and can use these services as an extension of its internal, potentially non-cloud services. It is possible for different cloud services to use different identity and credential providers, which are likely different from the providers used by the organization for its internal applications. The credential management system used by the organization must be consolidated or integrated with those used by the cloud services.

Authorization: Requirements for user profile and access control policy vary depending on whether the cloud user is a member of an organization, such as an enterprise, or as an individual. Access control requirements include establishing trusted user profile and policy information, using it to control access within the cloud service, and doing this in an auditable way. Once authentication is done, resources can be authorized locally within the CSP. Many of the authorization mechanisms that are used in traditional computing environments can be utilized in a cloud setting.

III LITERATURE SURVEY

Several mechanisms have been designed to allow both data owners and public verifiers to efficiently audit cloud data integrity without retrieving the entire data from the cloud server. However, public auditing on the integrity of shared data with these existing mechanisms will inevitably reveal confidential information—identity privacy—to public verifiers. Cloud service providers offer users efficient and scalable data storage services with a much lower marginal cost than traditional approaches [2]. It is routine for users to leverage cloud storage services to share data with others in a group, as data sharing becomes a standard feature in most cloud storage offerings, including Dropbox, iCloud and Google Drive.

Recently, many mechanisms [9]] have been proposed to allow not only a data owner itself but also a public verifier to efficiently perform integrity checking without downloading the entire data from the cloud, which is referred to as public auditing [10]. In these mechanisms, data is divided in to many small blocks, where each block is independently signed by the owner; and a random combination of all the blocks instead of the whole data is retrieved during integrity checking [11]. A public verifier could be a data user (e.g., researcher) who would like to utilize the owner's data via the cloud or a third-party auditor (TPA) who can provide expert integrity checking services [12]. Moving a step forward, [13] designed an advanced auditing mechanism [5], so that during public auditing on cloud data, the content of private data belonging to a personal user is not disclosed to any public verifiers. Unfortunately, current public auditing solutions mentioned above only focus on personal data in the cloud.

Provable data possession (PDP), proposed [13], allows a verifier to check the correctness of a client's data stored at un-trusted server. By utilizing RSA-based homomorphic authenticators and sampling strategies, the verifier is able to publicly audit the integrity of data without retrieving the entire data, which is referred to as public auditing. Unfortunately, their mechanism is only suitable for auditing the integrity of personal data. [14] Defined another similar model called Proofs of Retrievability (POR), which is also able to check the correctness of data on an un-trusted server. The original file is added with a set of randomly-valued check blocks called sentinels.

The verifier challenges the un-trusted server by specifying the positions of a collection of sentinels and asking the un-trusted server to return the associated sentinel values.

[15] Designed two improved schemes. The first scheme is built from BLS signatures, and the second one is based on pseudo-random functions. To support dynamic data,[15] presented an efficient PDP mechanism based on symmetric keys. This mechanism can support update and delete operations on data; however, insert operations are not available in this mechanism. Because it exploits symmetric keys to verify the integrity of data, it is not public verifiable and only provides a user with a limited number of verification requests.

In [13] Utilized Merkle Hash Tree and BLS signatures to support dynamic data in a public auditing mechanism. [14]

Introduced dynamic provable data possession (DPDP) by using authenticated dictionaries, which are based on rank information. [15] exploited the fragment structure to reduce the storage of signatures in their public auditing mechanism. In addition, they also used index hash tables to provide dynamic operations on data. The public mechanism proposed by[12] are able to preserve users' confidential data from a public verifier by using random masking's. In addition, to operate multiple auditing tasks from different users efficiently, they extended their mechanism to enable batch auditing by leveraging aggregate signatures [13].

In [13] leveraged homomorphic tokens to ensure the correctness of erasure codes-based data distributed on multiple servers. This mechanism is able not only to support dynamic data, but also to identify misbehaved servers. To minimize communication overhead in the phase of data repair,[14] also introduced a mechanism for auditing the correctness of data under the multi-server scenario, where these data are encoded by network coding instead of using erasure codes. More recently, [15] constructed an LT codes-based secure and reliable cloud storage mechanism. Compare to previous work [13], [14], this mechanism can avoid high decoding computation cost for data users and save computation resource for online data owners during data repair.

Oruta [6], a privacy-preserving public auditing mechanism for shared data in the cloud. We utilize ring signatures to construct homomorphic authenticator's, so that a public verifier is able to audit shared data integrity without retrieving the entire data, yet it cannot distinguish who is the signer on each block. To improve the efficiency of verifying multiple auditing tasks, [6] further extend our mechanism to support batch auditing propose a novel privacy-preserving mechanism that supports public auditing on shared data stored in the cloud. In particular, exploit ring signatures to compute verification metadata needed to audit the correctness of shared data. With our mechanism, the identity of the signer on each block in shared data is kept private from public verifiers, who are able to efficiently verify shared data integrity without retrieving the entire file. In addition, our mechanism is able to perform multiple auditing tasks simultaneously instead of verifying them one by one.

III. CONCLUSION

Cloud computing is a computing in which large groups of remote servers are networked to allow centralized data storage and online access to computer services or resources. Several mechanisms have been designed to allow both data owners and public verifiers to efficiently audit cloud data integrity without retrieving the entire data from the cloud server. However, public auditing on the integrity of shared data with these existing mechanisms will inevitably reveal confidential information—identity privacy—to public verifiers. The major issue with the cloud database is that it requires a very high level security. Data are not always safe when they are stored inside cloud providers. Since these data-centers may be located in any part of the world beyond the reach and control of users, there are multifarious security and privacy challenges that need to be understood and addressed. Privacy preservation is main security issue in public cloud. In this paper we discuss the cloud database security. We review the adaptive encryption scheme.

V. REFERENCES

[1] Peter Mell,TimothyGrance,"The NIST definition of cloud computing",http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[2] InduArora ,Dr.AnuGupta, "Cloud database: A paradigm shift in Databases", IJCI ,july 2012.

[3] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic,"Cloud computing and emerging it platforms: Vision, hype, andreality for delivering computing as the 5th utility," Future GenerationComput. Syst., vol. 25, no. 6, pp. 599–616, 2009.

[4] Suraj S. Gaikwad, Amar R. Buchade, "Survey on Securing data using Homomorphic Encryption in Cloud Computing.", International Journal of Computer Sciences and Engineering, Volume-04, Issue-01, Page No (17-21), Jan -2016

[5]Shivlal Mewada, Umesh Kumar Singh and Pradeep Sharma, "Security Enhancement in Cloud Computing (CC)", ISROSET-International Journal of Scientific Research in Computer Science and Engineering, Vol.-01, Issue-01, pp (31-37), Jan -Feb 2013.

[6] Boyang Wang, Baochun Li and Hui Li, Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud, IEEE Transactions on Cloud Computing, Vol. 2, No. 1, January-March 2014, Pp. 43-57

[7] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 598-610, 2007.

[8] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT '08), pp. 90- 107, 2008.

[9] C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," Proc. 16th ACM Conf. Computer and Comm. Security (CCS'09), pp. 213-222, 2009.

[10] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamic for Storage Security in Cloud Computing," Proc. 14th European Conf. Research in Computer Security (ESORICS'09), pp. 355-370, 2009.

[11] C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. 17th Int'l Workshop Quality ofService (IWQoS'09), pp. 1-9, 2009.

[12] B. Chen, R. Curtmola, G. Ateniese, and R. Burns, "Remote Data Checking for Network Coding-Based Distributed Storage Systems," Proc. ACM Workshop Cloud Computing Security Workshop (CCSW'10), pp. 31-42, 2010.

[13] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds," Proc. ACM Symp. Applied Computing (SAC'11), pp. 1550-1557, 2011.

[14] N. Cao, S. Yu, Z. Yang, W. Lou, and Y.T. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, 2012.

[15] C. Wang, S.S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-PreservingPublic Auditing for Secure Cloud Storage," IEEE Trans. Computers, vol. 62, no. 2, pp. 362-375, Feb. 2013.

