Volume 5, Issue 5, May 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

# Cloud Database Security: A Survey

**Anjali Nayak**
M.Tech Scholar
Computer Science Department
LNCTS, Bhopal, (M.P.) India

**Dr. Sadhna K Mishra**
Head
Computer Science Department
LNCTS, Bhopal, (M.P.) India

***Abstract- 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. This paper focuses on various security mechanisms that are provided in the enterprises and also discusses few of the common security mechanisms like authentication, authorization, encryption and access control.***

***Keywords-Cloud database, Security, Encryption, Authentication, DaaS***

## I. INTRODUCTION

Cloud computing can be defined as new computing that has focus on both industry and academia. Cloud computing is the result of evolution and adoption of existing technologies and paradigms. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. 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. This cloud model is composed of five essential characteristics, three service models, and four deployment models[2].Massive growth in digital data, changing data storage requirements, better broadband facilities and Cloud computing led to the emergence of cloud databases .Cloud Storage, Data as a service (DaaS) 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. Dropbox, iCloud etc. are popularcloud 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 isonestep 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 [4].

## II. CLOUD COMPUTING SERVICES

There are main three types of cloud computing service models-

**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.[2] 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.[2] 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)-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.[3]

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 [4].

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.[5] 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.

Security issues in cloud database:

**Scalability**-Cloud database should be scalable so that it can store as much data as possible when number of users in the cloud increases.

**Heterogeneity-**Cloud database should support all types of users i.e. users working on various platforms.

**Confidentiality-**Data should be readable to the legitimate clients only that is a proper encryption is done to the database so that no unauthorized user can access data.

These above mention are some of the common challenges that has to be faced while developing a cloud database.

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 to
many customers and is managed by a third party [11]. 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 [11]. 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.

### *Access control*

Access management 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. ADAPTIVE ENCRYPTION SCHEME

The [5] consider SQL-aware encryption algorithms that guarantee data confidentiality and allow the cloud database engine to execute SQL operations over encrypted data. As each algorithm supports a specific subset of SQL operators, we refer to the following encryption schemes.

*Random (Rand):*It is the most secure encryption because it does not reveal any information about the original plain value. It does not support any SQL operator, and it is used only for data retrieval.

*Deterministic (Det):* It deterministically encrypts data, so that equality of plaintext data is preserved. It supports the equality operator.

*Order Preserving Encryption (Ope) :* It preserves in the encrypted values the numerical order of the original unencrypted data. It supports the comparison SQL operators (=,<,<=,>,>=).

*Homomorphic Sum (Sum)*: It is homomorphic with respect to the sum operation, so that the multiplication of encrypted integers is equal to the sum of plaintext integers. It supports the sum operator between integer values.

*Search*: It supports equality check on full strings (i.e., the LIKE operator).

*Plain:* It does not encrypt data, but it is useful to support all SQL operators on non confidential data.

If each column of the database was encrypted with only one algorithm, then the database administrator would have to decide at design time which operations must be supported on each database column. However, this solution is impractical for scenarios in which the database workload changes over time.The proposed system in [5] supports adaptive encryption

for public cloud database services, where distributed and concurrent clients can issue direct SQL operations. By avoiding an architecture based on intermediate servers between the clients and the cloud database, the proposed solution guarantees the same level of scalability and availability of the cloud service. Fig.1 [5] shows a scheme of the proposed architecture where each client executes an encryption engine that manages encryption operations. This software module is accessed by external user applications through the encrypted database interface. The proposed architecture manages five
Types of information:

1. Plain data represent the tenant information;
2. Encrypted data are the encrypted version of the plain data, and are stored in the cloud database;
3. Plain metadata represent the additional information that is necessary to execute SQL operations on encrypted data;
4. Encrypted metadata are the encrypted version of the plain metadata, and are stored in the cloud database.
5. Master key is the encryption key of the encrypted metadata, and is known by legitimate clients. All data and metadata stored in the cloud database are encrypted. Any application running on a legitimate client can transparently issue SQL operations (e.g., SELECT, INSERT, UPDATE and DELETE) to the encrypted cloud database through the encrypted database interface. Data transferred between the user application and the encryption engines are not encrypted, whereas information is always encrypted before sending it to the cloud database. When an
Application issues a new SQL operation, the encrypted database interface contacts the encryption engine that retrieves the encrypted metadata and decrypts them with the master key.

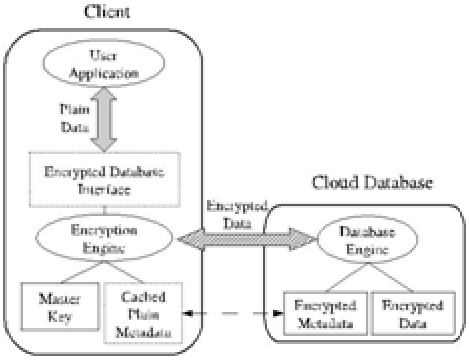

Figure 1 Encrypted cloud database

To improve performance, the plain metadata are cached locally by the client. After obtaining the metadata, the encryption engine is able to issue encrypted SQL statements to the cloud database, and then to decrypt the results. The results are returned to the user application through the encrypted database interface. As in related literature, the proposed architecture guarantees data confidentiality in a security model in which: the network is untrusted, tenant users are trusted, that is, they do not reveal information about plain data, plain metadata, and the master key; the cloud provider administrators are defined semi-honest or honest-but-curious, that is, they do not modify tenant's data and results of SQL operations, but they may access tenant's information stored in the cloud database. The remaining part of this section describes the adaptive encryption schemes, the encrypted metadata stored in the cloud database, and the main operations for the management of the encrypted cloud database.

Metadata concept used in the [5] is also considered in the proposed work. After generating the multi-user key its distribution is also done in the algorithm. A cost model can also be derived which will tell the mathematical representation of the algorithm performance. Java platform will be used for the implementation of the algorithm and MYSQL server is used as the back-end. Linux operating system is considered good in the security point of view thus except for windows OS Linux operating system will be used.

We are interested in the database as a service paradigm that poses several research challenges in terms of security and cost evaluation from tenant's point of view. Most results concerning encryption for cloud database services are inapplicable to the database paradigm. Other encryption schemes that allow the execution of SQL operations over encrypted data either have performance limits or require the choice of which encryption scheme must be adopted for each database column and SQL operation. These proposals are fine when the set of queries can be statically determined at design time, while we are interested in other common scenarios where the workload may change after the database design. In previous paper [5] a novel architecture for adaptive encryption of public cloud databases is proposed that offers a proxy-free alternative to the system described. This proposed architecture guarantees in an adaptive way the best level of data confidentiality for any database workload, even when the set of SQL queries dynamically changes. The adaptive encryption scheme, which was initially proposed for applications not referring to the cloud, encrypts each plain column to multiple encrypted columns, and each value is encapsulated in different layers of encryption, so that outer layers guarantee higher confidentiality but support fewer computation capabilities with respect to the inner layers. The outer layers are dynamically adapted at runtime when new SQL operations are added to the workload. In previous work [5], the first proxy-free architecture for adaptive encryption of cloud databases that does not limit the availability, elasticity and scalability of a plain cloud database because multiple clients can issue concurrent operations without passing through some centralized component as in alternative architectures. Although data encryption seems the most intuitive solution for confidentiality, its application to cloud databases services is not trivial, because the cloud database must be able to execute SQL operations directly over encrypted data without accessing any decryption key. Naive solutions encrypt the whole database with some standard encryption algorithm that does not allow executing any SQL operation directly on the cloud. As the consequence, the tenant has two alternatives: download the entire database, decrypt temporarily the cloud database, execute the query and, if the operation modifies the database, encrypt and upload the new data, decrypt temporarily the cloud database, execute the query, and re-encrypt it. The former solution is affected by huge communication and computation overheads, and consequent costs that would make cloud database services quite inconvenient, the later solution does not guarantee data confidentiality because the cloud provider obtains decryption keys. The right alternative is to execute SQL operations directly on the cloud database, without giving decryption keys to the provider. Many algorithms were proposed for this work like aggregation technique, fully homomorphism encryption etc. The drawback related to these feasible encryption algorithms is that in a medium-long term horizon, the database operations will be required over each database column. A solution to these problems were then given , the proposed architecture allows multiple clients to issue concurrent SQL operations to an encrypted database without any intermediate trusted server, but it assumes that the set of SQL operations do not change after the database design. The [5] develops the initial design through a prototype implementation, novel experimental results and an original cost model.

## IV. 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.The major issue with the clouddatabase is that it requires a very high level security.

## REFERENCES

[1] White paper cloud computing,"Alternative sourcing strategy for businessICT",T-syytementerpriceservice,http://tsystemsus.com/white_paper_cloud_computing/257d.pdf

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

[3] RajkumarBuyya,"Introduction to the IEEE transactions on cloud computing" vol 1,januarry-june 2013

[4] InduArora ,Dr.AnuGupta,"Clouddatabase:A paradigm shift in Databases"IJCI international journal,july 2012.

[5] Lucca Ferretti,FabioPierazzi,MichelColajani and Micro Marchetti "Performance and cost evaluation ofan adaptive encryption architecture for cloud databases"IEEE transactions on cloud computing,vol 2,no.2,April-June 2014

[6] MargerateRouse,searchcloudcomputing,http://searchcloudcomputing.techtarget.com/definition/cloud-computing

[7] MargerateRouse..http://searchcloudapplications.techtarget.com/definition/cloud-database-database-as-a-service

8] 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.

[9] T. Mather, S. Kumaraswamy, and S. Latif, Cloud Security and Privacy:An Enterprise Perspective on Risks and Compliance. Sebastopol,CA, USA: O'Reilly Media, Inc., 2009.