SURVEY ON ALGORITHM TO PERFORM SQL OPERATIONS OVER ENCRYPTED DATABASE

J. Angeline Martina[1], M. Mohan Kumar[2]

[1]PG Scholar, [2]Asst. Professor

Dept of Computer Science Engineering,

Sri Vidya College of Engineering and Technology (India)

ABSTRACT

The cloud computing has reap its prevalence by anything has a service (Xaas).The three major services are Software as a Service, Platform as a Service and Infrastructure as a service. The database is a maturing service nowadays. The database stored at the cloud should be encrypted and operations on the encrypted data to be performed inorder to reduce the passive attack. The passive attacks exist inside the system but it does not affect the data, sends information to the intruder. The Snooping of data can be reduced by performing SQL query operations over encrypted database. In accordance to encrypt and perform SQL query operations over database the various algorithms are used. This paper provides an overview of survey about various algorithms like homomorphic, AES in various modes are modified to support operations over the encrypted database.

Keywords— ***Cloud Computing, DBAAS, Encryption, Security, SQL Query Operations.***

I. INTRODUCTION

The cloud computing has gain its popularity from anything as a service. The major services offered by cloud are Software as a service, Platform as a service and Infrastructure as a service. The infrastructure as a service provides the storage for the users.[1][13]The Database-as-a-service provides an access to database without the need for hardware setup, installation and configuration. The user can use the database and pay as per they use. The administrative and maintenance are taken care by the providers. In essence, DbaaS is a managed service offering access to a database to be used with applications and their related data. Database-as-a-service is a structured approach based on Storage-as-a-service and it is similar to software as a service. The payment includes not only the usage but also for the administrative and maintenance of database provided by the database provider.[2] The database manager component controls all the instances of the database by an application program interface. The application program interface is accessible by web application and management console to the user and the user may manage, configure the database and allocate or deal locate the database instances. Although the emergence of cloud computing is a recent development, insights into critical aspects of security can be gleaned from reported experiences of early adopters and also from researchers analysing and experimenting with available cloud provider platforms and associated technologies. The organization's ownership rights over the data must be firmly established in the service contract to enable a basis for trust and privacy of data. With. The organisations have much equipment to secure the data which is vulnerable to attack in which; if the data is sensitive it is not shared. Risk Management is the process to identify and reduce the organizational risk assets, operations and individuals risk. The cloud is a public pool and data stored in cloud resides in a shared environment. As cloud is a shared pool of storage, it is vulnerable to attack so the access

control should be given by the data owner. The encryption and decryption is necessary for the cloud because the data can be shared by one user to another. The cipher text is an encrypted data which cannot be easily understand by the unauthorized users. The process of rolling back the cipher text to the plaintext text which can be understand is an decryption. The database stored in the cloud should be encrypted. Inorder to avoid the passive attack the data processing should be on the encrypted database. The operations can be performed by using the homomorphic encryption[9]. Homomorphic encryption would allow the chaining together of different services without exposing the data to each of those services. The homomorphic encryption allows the user to perform SQL operations on the encrypted database[10][11].The snooping of database is observing the information about the database stored in the cloud. The security also includes the three major features Confidentiality. Integrity and avalaibility.The passive attack affects the confidentiality while the active attack affects the integrity and availability.

II. DATABASE-AS-A-SERVICE

The physical architecture provides an understanding of the major technology components and their relationships in support of the DBaaS capabilities, processes and services. In the Logical Architecture, examples focused on DBaaS Management Capabilities. In this section, examples will focus on the fundamental DBaaS capabilities related to supporting the deployment of database instances. The deployment and management services must be implemented carefully to ensure that they are secure, reliable, scalable, and support all availability and service continuity requirements.[2]

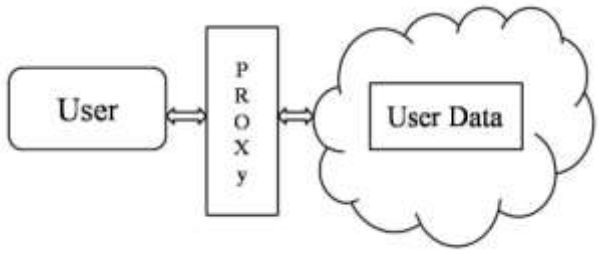

FIG.1 DBaaS Architecture

Advantages

1).The user need not to install the software

2).The DBaaS is cost effective.

Disadvantages

1).The User does not has the full control over the database.

2).The database as a Service is an sticky services.

III. EXISTING ARCHITECTURE

A novel architecture that integrates cloud database services with data confidentiality and the possibility of executing concurrent operations on encrypted data and it eliminates proxy server. The database administrator encrypts the database and dropped it onto the cloud. The client has the plain database with encrypted database and metadata. The metadata stored in the cloud are encrypted. The client can only have the access to the cloud. The client can request the required data from the database. [5]

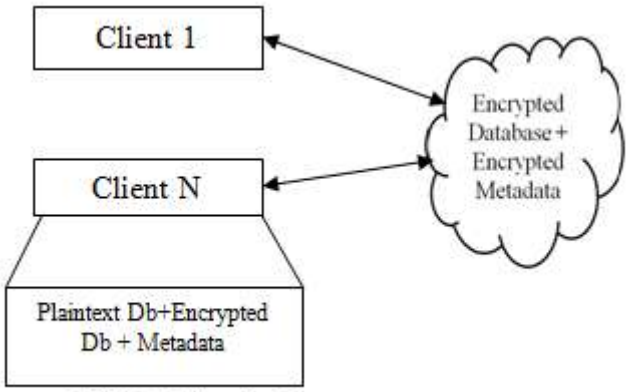

FIG.2 Existing Architecture

It integrates cloud database services with data confidentiality and the possibility of executing concurrent operations on encrypted data. This is the first solution supporting geographically distributed clients to connect directly to an encrypted cloud database, and to execute concurrent and independent operations including those modifying the database structure. The proposed architecture has the further advantage of eliminating intermediate proxies that limit the elasticity, availability, and scalability properties that are intrinsic in cloud-based solutions.

3.1 Description of Components

- **Client**

  The clients are the finite number of authorized user who can access data directly from the cloud. The clients are users with respect o the cloud.

- **User**

  The users are the authorized entity who can request the database and can access cloud by using the client.

- **Plaintext**

  The Plaintext is the unencrypted database which resides in the client

- **Encrypted Database**

  The encrypted database which resides both in the client and the cloud. The encryption algorithm used is AES in CBC mode.

- **Metadata**

  The metadata has the information about the plaintext and the encrypted database. The metadata is information which has plaintext table name, its encrypted table name and its field name.

- **Cloud**

  The cloud is used to store the encrypted database and the metadata.

IV. PROPOSED ARCHITECTURE

The approach is to execute queries over encrypted data and the various algorithms are used to efficiently perform SQL operations over encrypted database. The user send query to the query processor and the query processor converts the SQL query to encrypted query then forward to the cloud. The encrypted result is send back to the query processor and decrypted result to the user. The SQL query is converted by using the metadata. The coder and transmitter has the database, metadata the metadata is stored in the form of tree structure.

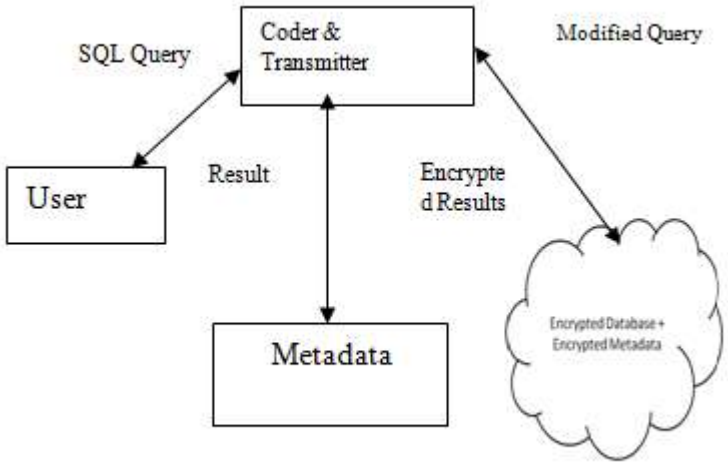

Fig.3 Proposed Architecture

4.1 Description of Components

- **User**

  The users are the authorized entity who can request the database and can access cloud by using the client. The user can send SQL command to the client.

- **Coder and Transmitter**

  The encoder and transmitter are the finite number of authorized user who can access data directly from the cloud .The encoder and transmitter rewrites the query with respect to encrypted database. The encoder and transmitter also decrypts the encrypted result and send the actual result to the user

- **Metadata**

  The metadata has the information about the plaintext and the encrypted database. The metadata is information which has plaintext table name, its encrypted table name and its field name. The metadata also holds the which type of encryption algorithm used for the records.

- **Cloud**

  The cloud holds the encrypted database and metadata.

V. SURVEY OF ENCRYPTION ALGORITHM

Database encryption is the process of converting data, within a database, in plain text format into a meaningless cipher text by means of a suitable algorithm.

Database decryption is converting the meaningless cipher text into the original information using keys generated by the encryption algorithms. Database encryption can be provided at the file or column level.

5.1 Advanced Encryption Standard

AES is a symmetric-key block cipher for secure and classified data encryption and decryption. The AES has three fixed 128-bit block ciphers with cryptographic key sizes of 128, 192 and 256 bits. Key size is unlimited, whereas the block size maximum is 256 bits. The AES design is based on a substitution-permutation network (SPN) and does not use the Data Encryption Standard (DES) Feistel network. The AES replaced the DES with new and updated features Block encryption implementation, 128-bit group encryption with 128, 192 and 256-bit

key lengths, Symmetric algorithm requiring only one encryption and decryption key. *S*ub Bytes is an 8-bit function. A look-up table that implements Sub Bytes contains 256 entries that are 8-bit wide. The Shift Rows operation changes the order of bytes in each row of the state. Mix Columns is a 32-bit operation that transforms four bytes of each column in the state. The key expansion routine is used to generate the round keys from the cipher key. The modes of operation are Electronic Codebook Mode (ECB), Cipher Block Chaining Mode (CBC), Cipher Feedback Mode (CFB), Output Feedback Mode (OFB), and Counter Mode (CTR).

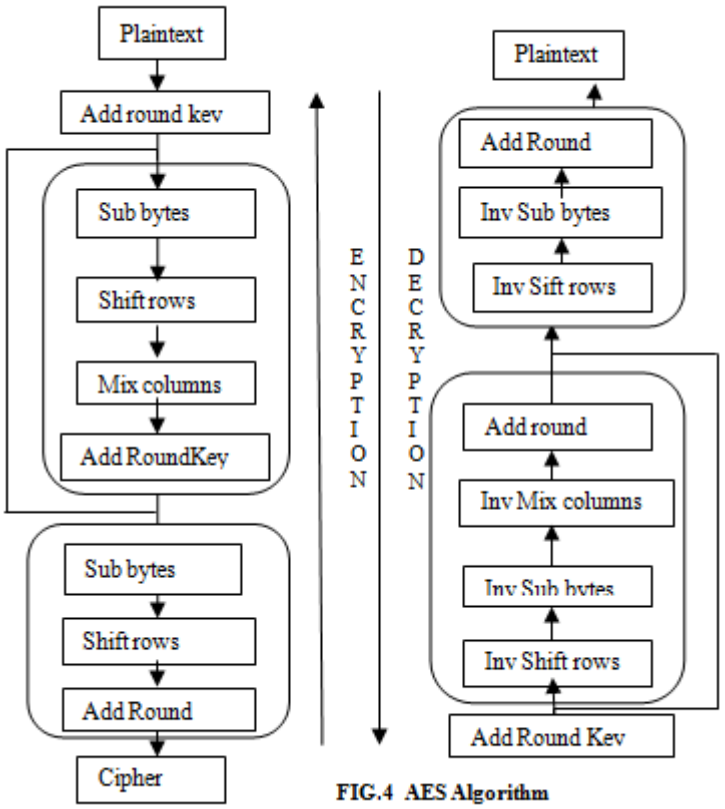

FIG.4 AES Algorithm

Advantages

1). AES is more secure.

2).AES is faster in both hardware and software.

3).Key size can be upto 256 bytes

Disadvantages

1). Too simple algebraic structure

5.2 Homomorphic encryption

Homomorphic encryption is the conversion of data into cipher text that can be analysed and worked with as if it were still in its original form.[15]Homomorphic encryptions allow complex mathematical operations to be performed on encrypted data without compromising the encryption. In mathematics, homomorphic describes the transformation of one data set into another while preserving relationships between elements in both sets. Homomorphic encryption is expected to play an important part in cloud computing, allowing companies to store encrypted data in a public cloud and take advantage of the cloud provider's analytic services. **The paillier algorithm** is used for the additive and product operations.[8][16]

Key Generation

1) P, q prime numbers

2) $n=p,q$, n^2

3)$\lambda=lcm(p-1,q-1)$

4)$\mu=(L(g^{\lambda} \bmod n^2))^{-1} \bmod n$

$L(u)=(u-1)/n;$

Public Encryption key **(n,g).**

Private Encryption Key **(λ,μ).**

Encryption

i)m,plaintext

ii)$r \varepsilon n^2$.

iii)$c = g^m.r^n \bmod n^{2.}$

Decryption

$m=L (c^{\lambda} \bmod n^2).\mu \bmod n$

Advantages

1). Paillier is a semantically secure, additively homomorphic cipher.

2).No need for the decryption of the data.

Disadvantages

1).Performance is the major concern because the ciphertext are larger than the plaintext.

5.3 Message Authentication Code

MAC involves two aspects Source authentication, which verifies the identity of the source, prevents the acceptance of messages from a fraudulent source. Data integrity, which protects the data from modification. Message authentication codes (MACs) attached to the message must be recognized by the receiving system in order to grant the user access. [20]

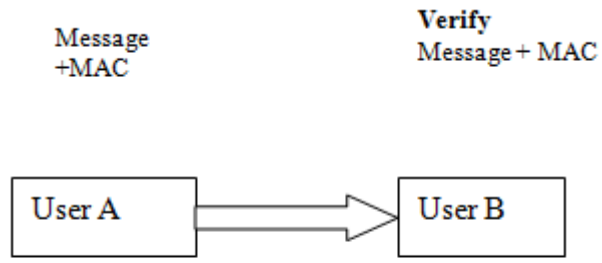

Fig.5 Message Authentication Code

One-time MAC Universal hashing and in particular pairwise independent hash functions provide a message authentication code as long as the key is used at most once**.** This can be seen as of the one-time pad for authentication **.**The one time message authentication is used to provide the integrity to the database stored at the cloud. [12]

Algorithm

Step: 1 The cipher text (Table Name) is written in the full alphabet. For example if ciphertext has 3 then write it as three.

Step: 2 The key is selected as equal to the length of ciphertext.

Step: 3 Then ciphertext, Key are Added then stored.

Step: 4 Then for Decryption the ciphertext are subtracted.

Advantages

1).Easy to compute.

2).Encryption and Decryption are same process

Disadvantages

1).Key must be long as the plaintext.

5.4 RC5 Algorithm

RC5 a parameterized algorithm with a variable block size, a variable key size, and a variable number of rounds. Allowable choices for the block size are 32 bits ,64 and 128 bits.

The number of rounds can range from 0 to 255, while the key can range from 0 bits to 2040 bits in size. RC5 has three routines: key expansion, encryption, and decryption.[14]

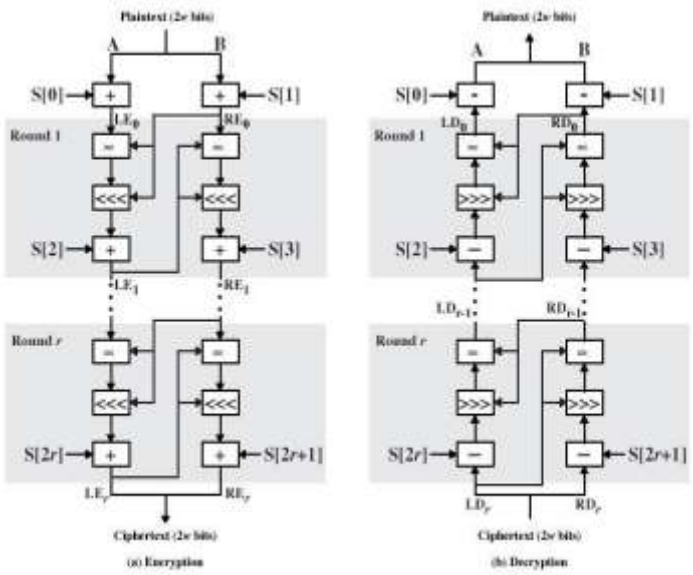

FIG. 6 RC5 Encryption

RC5 uses 3 primitive operations and their inverses:

- Addition modulo 2^w - +
- XOR -
- Left circular shift denoted by x<<<y, where word x is rotated y bits. Right circular shift is denoted by x>>>y.

5.5 Deterministic Algorithm

The Advanced Encryption Standard (AES) is a symmetric-key block cipher for secure and classified data encryption and decryption. The AES has three fixed 128-bit block ciphers with cryptographic key sizes of 128, 192 and 256 bits.

Key size is unlimited, whereas the block size maximum is 256 bits. The AES design is based on a substitution-permutation network (SPN) and does not use the Data Encryption Standard

(DES) Feistel network. The wide block encryption are used to securely encrypt the sectors of a disk called tweakable modes.[6]The CMC is an example used for the tweak able mode.[17]

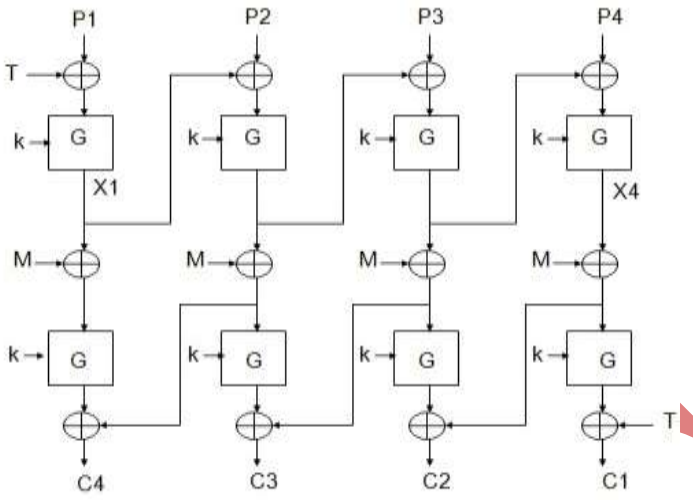

FIG.7 AES in CMC

VI. ENCRYPTION ALGORITHM

The various encryption algorithm are used to perform the SQL operations over encrypted database.

SCHEME	CONSTRUCTION	FUNCTION
RND	RC5	No Operation
HOM	PAILLIER	+,*
DET	AES IN CMC	Equality
INT	MAC	Verify

Table -1 Encryption Algorithm

With the required changes to the above algorithm the SQL operations are performed effectively. The database encrypted with a random algorithm where there is no need for the operations on the encrypted database. The paillier encryption which is the homomorphic encryption used for Sum and Count Operations. The deterministic encryption is used for Equality operations. The order preserving encryption used for aggregate functions. The Onetime MAC is used for the integrity verification [3][5].

VII. CONCLUSION

The database stored in the cloud is to be encrypted in order to provide the security.Inorder to reduce the passive attack the operations should be performed on the encrypted database. The Snooping of data can be reduced by performing SQL query operations over encrypted database. To encrypt and perform SQL query operations over database the various algorithms are used. The various algorithms are the random encryption, deterministic encryption, homomorphic encryption The various algorithms are discussed to perform the SQL operations over the encrypted database. To evaluate the performance of the system, time is measured by performing the operation over encrypted database. The Mac is used to verify the integrity.

REFERENCES

[1] M. Armbrust et al., "A View of Cloud Computing," Comm. of the ACM, vol. 53, no. 4, pp. 50-58, 2010.

[2] View Waleed Al Shehri "Cloud Database:Database as a Service" International Journal of Database Management Systems (IJDMS) Vol.5, No.2, April 2013.

[3] Divyakant Agrawal, Amr El Abbadi, Shiyuan Wang: Secure and Privacy-Preserving Data Services in the Cloud: A Data Centric View . PVLDB 5(12): 2028-2029 (2012)

[4] C. Curino, E. P. C. Jones, R. A. Popa, N. Malviya, E. Wu,S. Madden, H. Balakrishnan, and N. Zeldovich. Relational cloud: A database-as-a-service for the cloud. In Proceedings of the 5th Biennial Conference on Innovative Data Systems Research, pages 235–241, Pacific Grove, CA, January 2011.

[5] Luca Ferretti et.al.., "Distributed, Concurrent, and Independent Access to Encrypted Cloud Databases" IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 2, FEBRUARY 2014

[6] R.A. Popa, C.M.S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: Protecting Confidentiality with Encrypted Query Processing," Proc. 23rd ACM Symp. Operating Systems Principles, Oct. 2011.

[7] Sun S.Chung,et.al.., "Processing Aggregation Queries over Encrypted database."Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on April 2006

[8] Bharath K.Samanthula et.al.., "Privacy-Preserving Complex Query Evaluation over Semantically Secure Encrypted Data" 19th European Symposium on Research in Computer Security, Wroclaw, Poland, September 7-11, 2014. Proceedings, Part I Lecture Notes in Computer Science, Vol. 8712: Security and Cryptology

[9] Payal V.Parmar..et.al.., "Survey of Various Encryption algorithm and Schemes." International Journal of Computer Applications 2014 by IJCA Journal Vol 91 – Num 8,Year of Publication: 2014

[10] Bijiit Hore.et.al.., "Secure Multidimensional Range Queries over Outsourced Data" The VLDB Journal — The International Journal on Very Large Data Bases, Volume 21 Issue 3, June 2012, Pages 333-358

[11] Hui Wang..et.al.., "Efficient Secure Query Evaluation over Encrypted XML Databases" VLDB '06 Proceedings of the 32nd international conference on Very large data bases Pages 127-138

[12] Ajeet Ram..et.al.., "Survey of Confidentiality and Intergrity in Outsourced Databases."International Journal of Scientific Engineering and Technology,Volume 2 Issue 3, PP : 122-128

[13] http://en.wikipedia.org/wiki/Cloud_computing

[14] https://www.grc.com/r&d/rc5-report.pdf

[15] http://en.wikipedia.org/wiki/Homomorphic_encryption

[16] http://en.wikipedia.org/wiki/Paillier_cryptosystem

[17] http://en.wikipedia.org/wiki/Disk_encryption_theory

[18] http://crypto.stackexchange.com/questions/3813/how-does-order-preserving-encryption-work

[19] http://css.csail.mit.edu/cryptdb

[20] http://en.wikipedia.org/wiki/One-time_pad