SURVEY ON THE DATABASES IN THE CLOUD ENVIORNMENT

K. Priyatham[1], Arif Mohammad Abdul[2], M. Shanthi[3]

[1]Student, Department of Computer Science, GITAM University, Hyderabad
[2]Assistant Professor, Department of Computer Science, GITAM University, Hyderabad
[3]Assistant Professor, Department of Computer Science, GITAM University, Hyderabad

Abstract

The cloud database in present days plays a vital role in the organizations. The enterprises are adapting to the secured cloud services for storing their sensitive data in the database. Many organizations are adapting to different databases depending upon the requirement and the security provided by the database to the data. Many web applications are used by the organizations to increase the security provided by the database. These tools produce anomalous results in improving the security to the data in the database along with some demerits. This paper deals about the usage of the web applications which increase the level of security to the data in the cloud.

Keywords *– Cloud computing, Cloud database, SQL data model, NOSQL data model, Fully Homomorphic Encryption.*

----------***----------

1. INTRODUCTION

The organizations that are adapting the cloud services requires database to store their sensitive data. In order to ensure the confidentiality of the data in the database, different databases and tools are used. One of the major issues in the cloud services is security. So the organizations are not completely adapting the services of cloud.

Organizations are using various tools along with the different databases in order to ensure the security provided to the data in the cloud database. The DBA is the one who is responsible for providing the security to the data. So the DBA must gain the thrust of the organization to utilize the cloud database services.

Depending upon the usage and the security provided, there are several databases that are available with cloud services. Microsoft SQL Azure database, Xeround, SimpleDB, MongoDB, ClearDB etc. came into existence depending upon the requirement of the organization. As there are number of databases available, these are classified according to their data model as SQL data model, NOSQL data model etc.

Xeround, salesforce.com all these databases belongs to SQL data model while MongoDB, ClearDB are NOSQL data model. On the other hand, Microsoft SQL Azure database offers Azure SQL database as a standalone service. The NOSQL database is a simple database, where the SQL database requires SQL commands to perform any action on the database.

Though the database may be of any type, there are few threats to the databases. Attacking on the database by the unauthorized user or hacking the data and the compromising DBA are the major threats faced by the databases. As the sensitive data and the encryption keys are stored in the database, the DBA should not compromise to protect the security to the data. The other way for the theft of data is by using the methods like hacking, snooping etc without the knowledge of the client who is accessing the data in the database.

The above mentioned threats are the major threats where the organizations are trying to optimize and to maintain the confidentiality of the data in the database. Hence the organizations are using several tools along with the database to increase the level of security to the data. Though the tools are providing good result but they are not able to protect the data all the points. Some tools become weak during the transmission of data, or during the authentications of the multiple users etc which becomes a better source for grabbing the data by the unauthorized user.

2. LITERATURE SURVEY

The theft of sensitive data in the cloud of an organization is very often and it is the major factor for the organizations hesitating to adopt the cloud services completely. The best approach proposed in this paper is the usage of the servers like SUNDR and SPORC. SPORC provides a generic collaboration service in which users can create a document, update its access control list, edit it concurrently, experience fully automated merging of updates, and even perform these operations even though it is disconnected. The SPORC supports a broad range of collaborative applications.

The data which is sent by the client is stored in the MongoDB. Entire data is sent to the cloud database and then the data is encrypted and stored in the database. This MongoDB is also termed as NOSQL because it doesn't compute any SQL commands. MongoDB is a cross platform, which is document oriented database. This

database doesn't create number of files. Instead of taking a business subject and breaking it up into multiple relational structures, MongoDB can store the business subject in the minimal number of documents.

MongoDB supports search by index, regular expression searches etc. the queries given to the database returns specific fields of database and also includes some user-defined JavaScript functions. The MongoDB enables the feature of indexing any field. This helps during the searching of a particular data in the database. MongoDB can be used as a file system taking the advantage load balancing and data replication.

The usage of DAAS is very effective because of 2 reasons: (1) The amount of resource, man power, and the utilization of tools is not satisfactory because the amount invested is greater. (2) The usage of the concept "Pay per Use" which can be applied to both the software license and the administrative tools.

The process of encrypting the data takes place under the control of servers like SPORC, SUNDR etc. to provide confidentiality and to know the proper encryption of the data, which is to be sent to the database. This server doesn't store any data but notifies the client if any malicious server or any intruder attacks during the process of encryption of the data. Each different server has its own different methods for safeguarding the encryption of the sensitive data. These servers help in increasing the security to the process of encryption of data as the normal encryption doesn't notify about the intruder attacks and the malicious server during the encryption.

As the user updates the data in the database, the number of keys gets increased. This increase in number of keys results in slow down of the process and the problem of accessing the large number of keys. So the concept of multi key searchable encryption is used where, the data is encrypted by grouping the same data under a key. Therefore even the data increases, the growth of usage keys can be controlled and becomes easier for matching during the process of fetching or insertion of the data.

When the similar data is sent to the database then the same key is used to encrypt the file which decreases the use of large number of keys. This results in the easy storing of keys and the less usage of keys. Hence the processing of data becomes simpler even if it contains a large number of paired keys. In order to compute with the large number of keys, Lopez-Alt et algorithm is used. This algorithm results in best performance when compared with the other algorithms. This algorithm is designed with Fully Homomorphic Encryption (FHE) scheme where the users can evaluate the function for the data encrypted with different keys. A similar algorithm called Baoet algorithm is used for evaluating the multi-keys. In this algorithm, the users have a large number of keys where encryption and decryption of data done by using single key.

A special tool called Mylar is used as a platform to provide the confidentiality during the process searching the data in the cloud database. When the client sends a query or request for searching the data in the cloud, the data is encrypted at the client browser and the encrypted data is transmitted to the database. The encrypted query computes over encrypted data in the database and fetched the required encrypted data and it is sent to the client. This encrypted data is decrypted at the client's browser.

The Mylar tool has trusted hardware which can be trusted by the client. The other way for encryption of data is to rely on the hardware. The Mylar tools enable the secured encryption with the hardware depending upon the user trust worth. Mylar tool architecture has three parties namely, user or client, website owner and the server operator. Mylar provides security to the data during these three levels for ensuring the confidentiality of the user.

Hence the Mylar tool provides the special platform to the data search in the database. When the process of searching is performed on cloud database, then the secured data in becomes public and this data can be grabbed. The Mylar tool avoids this and remains the sensitive data as private data which results in security from the data hacking.

Cryptocat is used to provide security to the instant messages which are transferred between the users. This Cryptocat has threats, which are labeled as DREAD model.

The goal of Cryptocat is to increase the accessibility encrypted chat across platforms, languages and borders. In order to achieve this goal, the limits of Cryptocat system implementation in highly accessible environments must be expanded. Cryptocat mainly aims at the security to the instant messages without being hacked by any third parties.

The threats in Cryptocat are described as DREAD model which are defined as
D – Damage
R – Reproducibility
E – Exploitability
A – Affected Users
D – Discoverability

In Mylar both the client and the database servers is observed by the adversaries. So among the two treats for data database servers might compromise in security to the data easily. The adversaries might gain the data from client or from the database by using arbitrary responses. Mylar allows the client machines to control the adversaries to collude with the server. This is because the adversary may be the user or adversary might break into the client's machine. This model produces a wide range of security problems from bugs in server software to inside attacks.

3. SYSTEM MODEL

The purpose of SPORC is to provide the user to encrypt the data with confidentiality. SPORC enables the user to detect the malicious server in the network. The SPORC server

consists of list of authorized users which disables the unauthorized users to access the data of the authorized users. SPORC uses the Operational Transformation (OT) which provides general model for synchronizing the shared state. In OT, the application defines the modifications that can be applied to the document.

Though the SPORC is used for the security purpose, it has several threats. If the server is misbehaving or potentially malicious, then the server is able to prevent the progress but it should not affect the shared of the client, which is not happening under the server module. As the sever access to the size and timing of the user operations, it may clear the data. To avoid this threat, the data must be replicated with some other servers to provide the existence of the data.

In the client module, the presence of code authentication that can verify the part of code running on the client's machine is always assumed to be genuine. This process relies on the HTTPS connections to a trusted server.

SUNDR provides a different file system interface to remote storage, like NFS and other network file systems. As the file system is different, the location where the data is present remains safe and the intruder finds difficult to grab the data.

SUNDR uses a special protocol, which helps the authorized user to identify the modifications that attempted on the files by the unauthorized user in the network. SUNDR uses fetch-modify consistency, when the server is honest and acts perfect to the client and the fork consistency is used if the server is dishonest to the user. The fork consistency uses "time stamp box" to identify the actions that are performed with respect to the time.

SUNDR uses update certificates that are received from the client as a signed message. When the client sends a request then it is updated to the server by using the update certificates. If any update certificate is missed, then the server doesn't update the particular instruction to the data and hence results in failure in the process of encryption.

4. SYSTEM OVERVIEW

Mylar tool must concentrate in providing security to the data in all the three phases. So the Mylar tool encrypts the data each level to ensure the avoiding of data leakage at any level. Initially the data that is sent by the user is encrypted when it reaches web and it is sent to the database. The database again encrypts the encrypted file received from the client and it is stored in the database. The Mylar tool has two components namely, browser extension and client side library.

The browser extension is used for verifying the client code that is dumped by the server and ensures security in the client side application. The client side library intercepts the pair of keys that used by the client during the process of encryption and decryption. Client is also facilitated with the Identity Provider (IDP) which is used to authenticate the logged in client. This IDP is provided to the client by the website owner, which is used for generating the private key of the respective client by the website owner.

4.1 Security Overview

The main goal of Mylar tool is to provide the security to the client's sensitive data. Mylar tool checks the client side application and allows the client to perform the actions. The data sent by the client is encrypted at the user browser and it is sent to the database. This encrypted data is again encrypted by the database administrator and it is saved in the database.

The queries which are sent by the client are encrypted and sent to the database. The database administrator computes the query and gives the required file to the client in encrypted format. This file is again decrypted at the client's browser and viewed as original file to the client.

5. CONCLUSION

Among the number of tools in the cloud service, Mylar tool is one of the best tool which results in good performance with few drawbacks. This tool provides the user to search the data in the cloud database with good security to the sensitive data. The encryption and decryption of data takes place at different positions to increase the security provided to the sensitive data. The keyword search, which is not efficient in browser with general process, produces good result with the steps used by the tool. This tool increases the security to the data in the database during the process of searching the data in the cloud.

FUTURE WORK

During the process of interaction between the clients, the clients are not able to identify the unauthorized user when logged into the authorized users system. The limitations to the access of files in the database are not well known to the cloud database administrator which leads to improper usage of data. Redundancy of data is often found leading to wastage of cloud space. Sharing of files between the clients leads to the loss of confidentiality of the file as the user might get the file which the user doesn't have access to it.

REFERENCES

[1] Cipher Cloud. Cloud data protection solution http://www.ciphercloud.com.

[2] The concept of the MongoDB. http://en.wikipedia.org/wiki/MongoDB

[3] Cryptocat: Adopting Accessibility and Ease of Use as Security Properties by NadimKobeissi, Arlo Breault

[4] DAAS: Database As A Service by Carlo Curino, Evan P. C. Jones, Raluca Ada Popa, NirmeshMalviya, Eugene Wu, Sam Madden, HariBalakrishnan, NickolaiZeldovich

[5] Multi-Key Searchable Encryption by Raluca Ada Popa and NickolaiZeldovich

[6] CryptDB: Protecting Confidentiality with Encrypted Query Processing by Raluca Ada Popa, Catherine M.

S. Redfield, Nickolai Zeldovich, and Hari Balakrishnan MIT CSAIL

[7] SPORC: Group Collaboration using Untrusted Cloud Resources by Ariel J. Feldman.

[8] SUNDR: Secure Untrusted Data Repository (SUNDR) by Jinyuan Li, Maxwell Krohn, David Mazi`eres, and Dennis Shasha.

[9] The Cloud databases: http://en.wikipedia.org/wiki/Cloud_database

[10] Cloud database types: http://readwrite.com/2011/01/12/7-cloud-based-database-service.

[11] Building web applications on top of encrypted data using Mylar by Raluca Ada Popa, Emily Stark, Jonas Helfer, Steven Valdez, Nickolai Zeldovich, M. Frans Kaashoek, and Hari Balakrishnan MIT CSAIL and Meteor Development Group.

[12] Database management in Cloud Computing: http://www.itmanagerdaily.com/database-management/

[13] Featured databases in cloud computing: http://www.databasejournal.com/features/mssql/should-you-move-your-mysql-database-to-the-cloud.html

[14] Tools for providing security to cloud: http://www.hongkiat.com/blog/cloud-security-tools/

[15] The Cloud management: http://searchcloudcomputing.techtarget.com/report/Cloud-management-tools-guide-for-beginners

[16] Cloud computing tools: Improving security through visibility and automation: http://www.csoonline.com/article/2131715/identity-access/cloud-computing-tools--improving-security-through-visibility-and-automation.html