International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 05 | May 2019 www.irjet.net p-ISSN: 2395-0072

SECURING DATA FROM CSP WITH MULTICLOUD STRATEGY HANDLING NUMERIC RANGE QUERIES

Sinda P Xavier[1],

Student,Thejus Engineering College,Kerala,India

Abstract – *Cloud computing is the trend of the present world. Due to the tremendous features offered by the cloud infrastructure, many data owners are utilizing or exploiting the cloud for storing their personal data. Even they are exploiting the cloud; the data owners are bothered about their stored data. It is mainly because of the lack of security to the stored data. The main reason for this lack of security is the malpractice of the cloud service providers. The cloud service providers are considered to be the honest one but they are curious about the data which is stored by the data owners. So effective methods are necessary to protect the stored data from Honest but Curious cloud service providers. Here introduces a method for the securing the data from CSP by splitting the cloud service providers into two. That is explained as the multi-cloud strategy. Here the data which is stored in first CSP after the encryption process and the key used for the encryption is stored in second CSP. A new partitioned query processing strategy is also discussed here.*

Key Words: **Cloud computing, Honest but curious CSP, Multi-cloud Strategy, Partitioned query processing, Protection to stored data**.

1.INTRODUCTION

Cloud computing is the existing paradigm which is currently used by many data users. The features of cloud computing are widely exploited. The users which are exploiting the cloud infrastructure are increasing day by day. The storage and computation outsourcing is one of the popular exploiting feature of the cloud infrastructure. This helps the users to decrease the burden of the users in maintaining their personal data. Even though there are tremendous features are offered by the cloud system ; still there are some privacy concerns are related with the cloud infrastructure. This is mainly due to the insecure management of the cloud service providers. Usually the cloud system and the data stored in the cloud are under the control of the cloud manager called CSP. But actually the management by the CSP is under some security problems. The cloud service providers are considered as honest to a certain extent. But they are very much curious about the data stored in the cloud. That is why they are called Honest But Curious CSPs. So introduction to some methods to solve the security issues related with the CSPs are essential.

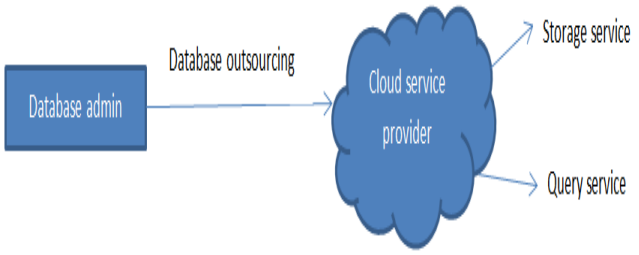


Fig -1: Database outsourcing

Generally there are so many methods are introduced to protect their data from cloud. But all these methods are usually operated among a single cloud. This recreates so many loop holes in the system to drop the data. This giving chances to the CSP to misuse the data which is stored by other users. CSPs are usually done this malpractice for business profits or something like that. Here introduces a new way to split the data storing method among two clouds. Here the database is stored in one cloud after encrypting the data. The key which is used for encrypting the cloud is uploaded to the second cloud. So the chances for getting pure data to CSP is less. The another method introduced here is to hide the query pattern from the CSP. So also splits the query processing logic among these two clouds. So through this way, the data contents as well as the query pattern can not be revealed to the CSP.

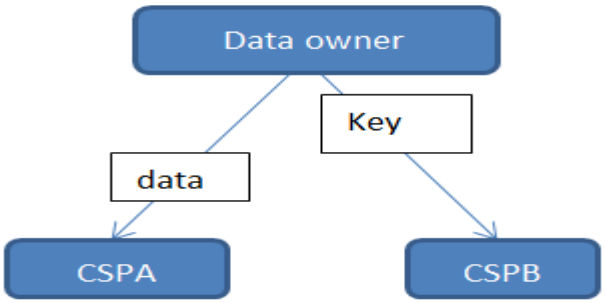
2. LITERATURE SURVEY

[1]Order preserving encryption is one of the most popular technique used for encrypting the numerical data before it is uploading to the cloud. This scheme is generally considered as a deterministic scheme to encrypt the data. But usually this method stores the numerical ordering of the data. This usually allows the CSP to get a guessed order of the way of encrypting the data. Similarly it also reveals the statistical properties of the queries. These loopholes are enough for the CSP to get the required data. So the order preserving encryption is not considered to be an effective method for encryption.[2]Multi-cloud architecture is introduced here so as to overcome the problems while operating with one single cloud. But here no suggestions are provided for the query processing logic. Here only explains the way to store the data securely. The encrypted database is stored to one cloud. The key which is used for

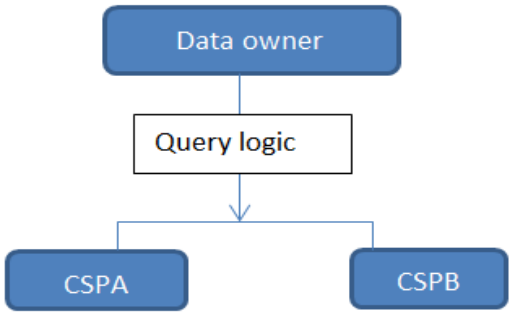
the encryption is stored in the second cloud. This basic idea is given by this paper. There is no clarity for the logic to process an query.[3]Fuzzy query are one type of queries which is operated over the encrypted data. These are mainly applicable in some practical scenarios. But still there are some practical issues are arises while applying the query operation. Fuzzy searchable encryption is a popular technique which is discussed in many papers. while applying the fuzzy queries ; generally numerical keywords are retrieved. This numerical keywords possess some distinction in character level. [4]CryptDB is another literature which referred to get an idea about the encryption mechanism on numerical range queries. CryptDB generally performs the operation on standard set of SQL queries. It usually applied for the preservation of confidentiality of the data contents. Generally, different type of encryption mechanism can be described on the CryptDB such as layered as well as adjustable encryption schemes.[5]KNN query scheme is another type of query scheme which is used for the query operations. It is generally used among the large relational databases. Usually the KNN query scheme is very simple. It works based on finding the distance between the sample points and query instance. But this seems to be expensive and it also reveals some information to a particular extent.[6]Another efficient way of query processing is to establish approximate query processing algorithm. There are so many methods are adopted to promote the query processing on online analytical processing. But due to one or another reason, no one is effective for the online query processing. At last arrived at an solution called approximate query processing. But actually this does not promotes the numeric range queries. That is why this query processing model is not adopted here.

3. METHODOLOGY

Generally, there are so many methods are adopted for securing the data from the cloud service providers. In most of the methods, a single cloud is used to store the entire data. Not only the data, but also the secret keys used for the encryption are also stored in the same cloud system. This paves the way for the cloud service providers to get the original data. Similarly the access to the cloud through queries will reveal the structure or pattern of the repeated queries. This is the main drawback which seen in the processing with a single cloud system. When using a single cloud system ; the cloud service providers can get knowledge about both the encrypted data and the way of encrypting them. So partitioning the knowledge among two clouds is the most acceptable way to avoid this criteria.



a) Data Partitioning



b) Query logic Partitioning

Fig -2: Multi-cloud architecture

So here adopted a new strategy to split the single cloud system to two separate cloud systems. They are specifically called as CSPA and CSPB. Generally the partitioning includes the partitioning of the data and the query processing logic. Here the data owner encrypting his data using an encryption algorithm. This encrypted database is uploaded to the first cloud. The keys used for encryption are given to the second cloud through a secure channel. And the keys will store securely in the second cloud. This is how generally the security problems related with the cloud service providers are avoided. The methodology section is divided as separate modules such as cloud setup module, query setup module, query processing module.

3.1 Cloud setup

The cloud setup phase is building the two clouds from a single cloud infrastructure. Here the data owner first of all rents the database for outsourcing the data into cloud. Before outsourcing the data, the client will encrypt the required data with some symmetric encryption algorithms. Usually, RSA or DES is generally used for this. After encrypting the data, the encrypted database will

outsourced to the first cloud. The public keys will also be attached with the encrypted database. Similarly the secret keys will passed to the second cloud. By knowing any one of the knowledge, the cloud service provider cannot get any idea about the outsourced data. This is clearly done based on the assumption that there is no collision among the two cloud systems.

3.2 Query setup

As per the multi-cloud architecture, the system gives importance to the numerical related data. So the queries are defined as numerical range queries. Numerical range queries contains the operators such as "<",">", "BETWEEN" one column. Combinations over one or more columns can also depicted using these type of queries.

3.3 Query processing

The query processing logic is discussed here. First of all the data owner encrypts his database and outsource it to the cloud database of the first cloud. The key which is used for the encryption is securely uploaded to the second cloud. So the data as well as the keys are securely stored in the cloud A and cloud B respectively. When a query predicates comes from a data user, the cloud A will receives the query and will converts the query from plaintext mode to cipher text mode. After taking the cipher text query, the first cloud will extract the column specified in the query. This column will undergo an item shuffling method and totally the whole items will shuffled. Then this shuffled items will pass on to the second cloud after removing the column name. This particular column will received at the second cloud through a secure channel. Using the secret key, the column get decrypted. The query operation will applied on this particular decrypted column and found out the indices of the data satisfying the query. These indices will pass on to the first cloud and here the data will trace out as per the indices. This data will pass on to the data owner and he will figure out the original data.

4. CONCLUSION

The method proposed here is a multi-cloud strategy to secure the data from the honest but curious cloud service providers. This paves the way to hide the data from the CSP. First cloud only stores the encrypted database. So there for he have no idea about the secret keys used for encrypting the database. Similarly the second cloud knows the key ; but he have no access to the encrypted database. So with the knowledge of any one of the parts, none of the cloud get an access to the original data and thereby it avoids the malpractices of the cloud service providers. The query processing logic is also an effective way to hide the query access patterns. Not only the data logic but also the query logic is partitioned among the cloud systems.so through this proposed system, the data contents, query patterns are safe from the honest but curious cloud service providers.

REFERENCES

[1] A. Boldyreva, N. Chenette, and A. O'Neill. "Order-preserving encryption revisited: improved security analysis and alternative solutions", in CRYPTO, 2011.

[2] D. Boneh, D. Gupta, I. Mironov, and A. Sahai, "Security and privacy enhancing multicloud architectures", in Advances in Cryptology-EUROCRYPT 2015 .Springer, 2013, pp. 404–436.

[3] M. Goncalves and L. Tineo. "SQLf Flexible Querying Language Extension by means of the norm SQLT", The loth IEEE International Conference on Fuzzy Systems, Vol 1, Dec 2001.

[4] R. A. Popa, C. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: protecting confidentiality with encrypted query processing," 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks,2017,pp. 234-241.

[5] R. A. Popa, C. Redfield, N. Zeldovich, and H. Balakrishnan, "Secure KNN query over encrypted data in outsourced environments" in Proceedings of the 23rd ACM Symposium on Operating Systems Principles.ACM, 2014, pp. 85–100.

[6] S. Benabbas, R. Gennaro, and Y. Vahlis, "Efficient SQL adaptive query processing in cloud architectures," in Annual Cryptology Conference.Springer, 2016, pp. 111–131.