Database Audit over Cloud Environment Using Forensic Analysis Algorithm

K. Govinda [#1], Pratik Nelge [*2], Mahesh Malwade.[#3]
[#] SCSE, VIT University
Vellore, India
[1] kgovinda@vit.ac.in
[3] malwade.mahesh@gmail.com
[*] SCSE, VIT University
Vellore, India
[2] prtknelge5@gmail.com

***Abstract*—Cloud Computing is a technology that uses the internet and virtual remote servers to maintain data and applications. Cloud computing allows consumers and organizations to use applications without installation and access their personal files at anytime and anywhere through internet access. It provides people the way to share distributed resources and services that belong to different organizations. Rapid growth in the field of cloud computing also increases severe security concerns due to distributed environment. When any unauthorized person tampers the data in database over cloud then data theft occurs. Such kind of fraud must be detected and necessary steps must be taken. In this paper we propose forensic algorithm to detect when tampering occurred and what data is tampered in the cloud database and propose method that gives more feasible solution.**

Keyword- Database Management, Security, Protection, Integrity, Secure master database

I. INTRODUCTION

Secure storage of data is need of every organization, businesses, banking sector etc. If data were to be changed illegally by whoever it may be, it might cause many harmful results for that organization as well as for their clients. There are many reasons for tampering the data such a student can change his grades or any person can change his bank account information such as amount etc. these are the outsider threads but there might be some insiders who can tamper the data illegally such as employee of company might change the data for his personal benefits. Practically most of the tampering done by the insiders than that of someone from outside[12].

Data outsourcing is the most growing field that allows user and organizations to give their data to external vendors who are responsible for storage and managing that data. This helps the organizations to concentrate only on their core business rather than management and storage of data. This also reduces their cost required for maintaining hardware, storing the data and maintaining that. Though outsourcing has many benefits it also introduces many concerns about security. This brings many challenges in data security as the organizations put their private data on the many shared servers which are not under the control of data owner. Third party i.e. the provider of that storage has the overall responsibility of the storing, managing and securing the data.It is proposed that limiting the access of information is not the proper way to secure the data. This proposal gives the concept of 'Information accountability' [10]. Information accountability has more advantage than the concept of restricting the information from the user. Information accountability successfully used from many years [11].Health Insurance Portability and Accountability Act-HIPAA [8] is relevant for privacy and security of health data and U.S. Public Law known as Sarbanes–Oxley Act [9] is relevant to certify security and accuracy of financial data.

II. LITERATURE REVIEW

This paper is extension of the work by Christian Collberg, Richard T. Snodgrass, and Shilong Stanley Yao. Multiple papers have been published which describes the ideas about detecting the tampering. Such as 'tamper detection in audit logs [4]' described in section A. We will discuss about the analysis of the database tampering [7] in the section B and also we will discuss the DRAGOON [5] system based on this concept and some forensic analysis algorithms in section C, .In these two (A and B) sections we will summarize these ideas, firstly describes how to detect tampering within a database and then secondly, how to analyse such tampering. The detailed literature survey explained in following sections. An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

A. Tamper Detection in Audit Logs

To prevent from illegal tampering the data by any intruder, it may an insider or outsider, many provisions are provided in the DBMS itself that uses cryptographically strong one-way hash functions. The DBMS periodically store the hash values of the data when it is committed in the database, it also stores the audit logs with the

transaction time. It also stores some information in database that enables the validator to check the database for its consistency. It is also noted that the database should append only. We cannot delete the records from the database instead we can update or add into the database. Past data should be regained. When the tuple is modified DBMS obtains the timestamp and its compute hash value using cryptographically strong one-way hash function and sends it to the notarizer. Then notarizer compute notary id for that hash value. It should also noted that calculating the notary id for each transaction is not possible so notarization should be done once per day when the hash values of transactions which are made during last 24 hours were hashed and its notary id's are calculated[4]. Validator is responsible for scanning audit tables, computing the hash values for each transaction and sends it to the digital notarization service along with its ID. It then reports if the data is modified or not. Suppose the intruder changes the data then hash values calculated by the validator would not be same with the hash values present in the audit logs. Auditing overhead was so negligible so that you can afford it for protecting your highly sensitive data.

B. Analysis of Database Tampering

The above concept explained in section 3.1 tells whether the data is corrupted or not. But this is not sufficient, further actions are to be taken to find out when the tampering has occurred, and actually what data has changed. Here we explain some terms as,A corruption event (CE) is the event that corrupts the data. When any intruder alters the data CE occurs. A notarization event (NE) is the notarization of hash value by the digital notarization service this event occurs as notarizer runs. A validation event (VE), it occurs when the validator runs. DBA schedules this validation event, so after fixed time interval this event occurs. When this event fails then tampering is detected.The forensic analysis detects both the parameters of tampering that is *when* the data was corrupted and at *which* location in the database it was corrupted using the corruption diagram [7]

C. DRAGOON system and Tiled Bitmap Forensic Analysis Algorithm

Using the above concepts Kyriacos E. Pavlou and Richard T. Snodgrass proposed a system called DRAGOON for detecting the tampering in the high performance databases [5]. In that system they used concept of storing the audit logs somewhere at the secure site [2].Here for forensic analysis some algorithms are explained such as A3D, monochromatic, RGB, RGBY [7], Tiled Bitmap. Amongst all above named algorithms Tiled Bitmap algorithm gives the more feasible solution. In this algorithm computed hash values are combined and hash chains are created [1]. By comparing these partially computed hash chains for each tile a binary string is generated which is given as input to the algorithm. Then the algorithm computes the candidate set. Candidate set represents all the possible combinations of where the tampering may be occurred. This is drawback of this algorithm that it generates false positive results. So we proposed a system, explained in the section 4, which extends the DRAGOON [5] to the cloud database and possibly removes the false positives from the result.

III. PROPOSED METHODOLOGY FOR TAMPER DETECTION

We are here proposing the architecture for detecting the tampering in the cloud database using above techniques. As we have seen that the tiled bitmap forensic analysis algorithm is the most optimal algorithm but still it gives some false positives. So in the following architecture we have done the analysis on how we can remove those false positives from the result.Here in the diagram we have shown two clouds because the working of system should be at some secure place which must be isolated from the main database. So cloud A contains the monitored database and the cloud B contains the secured storage. The overall working of the system is as follows.

When user application performs commits on the monitored database the hash values of that committed data are generated. Hash value generator is responsible for generating the hash values of the data. That generated hash values from the different transactions are linked with each other in order to generate the linked hashed chain which at each time instant represents the total data in the database. Before the calculated hash value is linked, it is stored with the timestamp in the separate database named as DB1. And the chains are stored into the secured database.

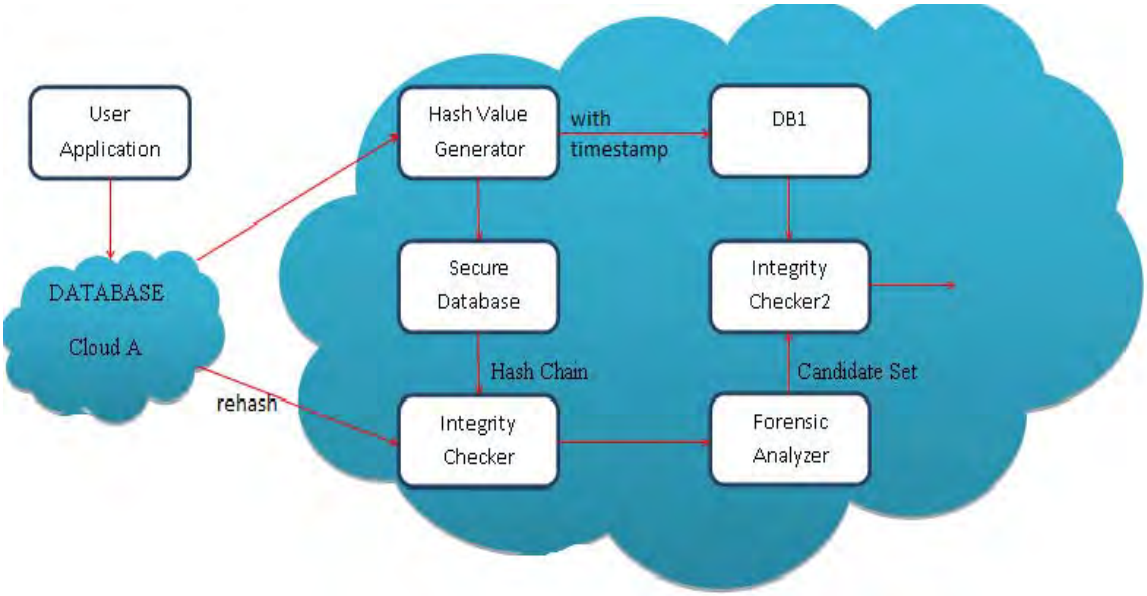

Figure1. System Architecture

The above Figure-1 also shows the tampering detection. After specific time period when the DBA initiates the validation of database for finding the tampering, the integrity checker came into picture. Then it again initiates the scan of entire database and computes the hash values over the database with the timestamp. Then it fetches the previously computed hash values from the secured database and compare with it. If values are not matched then integrity checker declares that the tampering is detected and it calls the forensic analyzer. It sends the binary string input to the algorithm computed from the comparing the hash chains. Forensic analyzer then using the Tiled bitmap algorithm finds the possible time values where the tampering might happened which is called as the candidate set. Then hash values of that time instant, which are the output of the analyzer, compared with the hash values that are stored in the DB1. Integrity checker2 performs this function, it checks that the values are inconsistent for which time instant. From this we get the exact time instant at which the tampering is happened.

IV. CONCLUSION

We explored the audit system for databases using forensic analysis method, and analyzed how it has been implemented in the DRAGOON database system. Security requirements are increasing as the government sectors and private organizations are now relying on the cloud The proposed method is not for just protecting the data and it continuously monitors and validates where and what data is tampered.This system has more feasible than the previous one which uses information restriction concept for avoiding the frauds in the database. The system is proposed in such a way that it will be easy for cloud vendors to implement in the cloud environment. In this way this system guarantees about the security of data which is the most essential requirement of the cloud vendors. As the system works efficiently for detecting the tampering, it gives little overhead on system to maintain extra database to store hash values, So further work needs to be done for reducing the overload from the system.

REFERENCES

[1] K. E. Pavlou and R. T. Snodgrass, “The tiled bitmap forensic analysis algorithm,” IEEE Trans. Knowledge Data Eng., vol. 22, no. 4, pp. 590– 601, April 2010.

[2] M. Malmgren, “An Infrastructure for Database Tamper Detection and Forensic Analysis,” Honors Thesis, University of Arizona, May 2007. http://www.cs.arizona.edu/projects/tau/tbdb/MelindaMalmgrenThesis.pdf.

[3] K. E. Pavlou and R. T. Sondgrass, “Achieving Database Information Accountability in the Cloud”, Data Engineering Workshops (ICDEW), 2012IEEE 28th International Conference on Digital Object Identifier: 10.1109/ICDEW.2012.37, Page(s): 147 – 150, 2012.

[4] R. T. Snodgrass, S. S. Yao, and C. Collberg, “Tamper Detection in Audit Logs,” in Proceedings of the International Conference on Very Large Databases, pp. 504–515, Toronto, Canada, September 2004.

[5] K. E. Pavlou and R. T. Sondgrass,” Dragoon an Information Accountability system for High Performance Databases”. Data Engineering (ICDE), 2012 IEEE 28th International Conference on Digital Object Identifier: 10.1109/ICDE.2012.139, Page(s): 1329 – 1332, 2012.

[6] HarmeetKaurKhanuja and D .S. Adane, ”Database Security Threats and Challenges in Database Forensic: A Survey”, International Conference on Advancements in Information Technology With workshop of ICBMG 2011, IPCSIT vol.20 (2011) ©(2011) IACSIT Press, Singapore.

[7] K. E. Pavlou and R. T. Snodgrass, “Forensic Analysis of Database Tampering,” in Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 109– 120, Chicago, June, 2006.

[8] U.S. Dept. of Health & Human Services, The Health Insurance Portability and Accountability Act (HIPAA), 1996, http://www.cms.hhs.gov/HIPAAGenInfo.

[9] U.S. Public Law No. 107-204, 116 Stat. 745. The Public Company Accounting Reform and Investor Protection Act, 2002.
[10] D. J. Weitzner, H. Abelson, T. Berners-Lee, J. Feigenbaum, J. Hendler, and G J.Sussman ,"Informationaccountability", *Communications of theACM*, vol. 51, no. 6,pp. 82–87,June 2008.
[11] R. G. Johnston, "Tamper-indicating seals," American Scientist, vol. 94, no. 6, pp. 515–524, Nov–Dec 2006.
[12] CSI/FBI, "Tenth Annual Computer Crime and Security Survey," July 2005,http://www.cpppe.umd.edu/ Bookstore/Documents/ 2005 CSISurvey.pdf.