Distributed Accountability for Data Sharing in Cloud

Snehal Suryawanshi
PG student of Department of Information Technology.
Pune Institute of Computer Technology
Sr. No 27, Dhankawadi, Pune Satara Road, Pune – 411043

Anant M. Bagade
Associate Professor Department of Information Technology
Pune Institute of Computer Technology
Sr. No 27, Dhankawadi, Pune Satara Road, Pune – 411043

ABSTRACT
Cloud computing is a technology, which uses internet and remote servers to stored data and application. Cloud computing provides on demand services. Multiple users want to do business of their data using cloud but they get fear to losing their data. While data owner will store his/her data on cloud , he must get confirmation that his/her data is safe on cloud. To solve above problem in this paper we provide effective mechanism to track usage of data using accountability. Accountability is checking of authorization policies and it is important for transparent data access. We provide automatic logging mechanisms using JAR programming which improves security and privacy of data in cloud. Using this mechanism data owner may know his/her data is handled as per his requirement or service level agreement**.**

General Terms
Cloud computing, accountability, security

Keywords
Cloud computing, accountability, security, data sharing, privacy

1. INTRODUCTION
Cloud computing is a technology which uses internet and remote servers to store data and application. In cloud there is no need to install particular hardware, software on user machine, so user can get the required infrastructure on his machine in cheap charges/rates. Cloud computing is an infrastructure which provides useful, on demand network services to use various resources with less effort. Features of Cloud computing are, huge access of data, application, resources and hardware without installation of any software, user can access the data from any machine or any where in the world, business can get resource in one place, that's means cloud computing provides scalability in on demand services to the business users. Everyone kept their data in cloud, as everyone kept their data in cloud so it becomes public so security issue increases towards private data. Data usage in cloud is very large by users and businesses, so data security in cloud is very important issue to solve. Many users want to do business of his data through cloud, but users may not know the machines which actually process and host their data. While enjoying the convenience brought by this new technology, users also start worrying about losing control of their own data [1], [8].
Cloud provides three service models, which are; platform as a service, infrastructure as a service and software as a service. Under the Database as a service, this is having four parts which are as per mentioned below,

- Encryption and Decryption - For security purpose of data stored in cloud, encryption seems to be perfect security solution.
- Key Management_- If encryption is necessary to store data in the cloud, encryption keys can't be store their, so user requires key management.
- Authentication - For accessing stored data in cloud by authorized users.
- Authorization – Rights given to user as well as cloud provider.

To solve the security issues in cloud; other user can't read the respective users data without having access. Data owner should not bother about his data, and should not get fear about damage of his data by hacker; there is need of security mechanism which will track usage of data in the cloud. Accountability is necessary for monitoring data usage, in this all actions of users like sending of file are cryptographically linked to the server, that performs them and server maintain secured record of all the actions of past and server can use the past records to know the correctness of action. It also provides reliable information about usage of data and it observes all the records, so it helps in make trust, relationship and reputation. So accountability is for verification of authentication and authorization. It is powerful tool to check the authorization policies [9].Accountability describes authorization requirement for data usage policies. Accountability mechanisms, which rely on after the fact verification, are an attractive means to enforce authorization policies [7].

There are 7 phases of accountability
1. Policy setting with data
2. Use of data by users
3. Logging
4. Merge logs
5. Error correctness in log
6. Auditing
7. Rectify and improvement.

These phases may change as per framework

First the data owner will set the policies with data and send it to cloud service provider (CSP), data will be use by users and logs of each record will be created, then log will be merged and error correction in log has been done and in auditing logs are checked and in last phase improvement has been done [12].

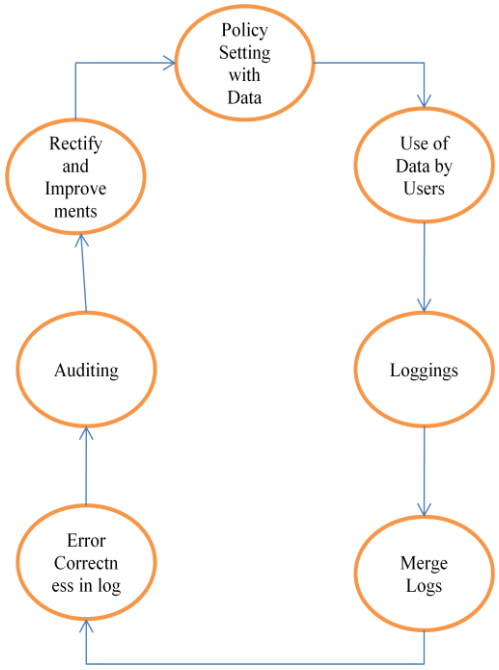

Fig 1: Phases of Accountability

In the Fig 1 Steps of accountability is given these are 7 steps each step is important to perform next step, accountability is nothing but validation of user actions means user having rights for accessing this data or not. Suppose user will do misuse of data or resources then network or data owner will take action on it so users, businesses and government should not bother about their data on cloud.

2. LITERATURE SURVEY

In this section review related works addressing security in cloud. Security issue is very important in cloud there are many techniques available so here is review of all these.

S. Pearson et al describes privacy manager mechanism in which user's data is safe on cloud , in this technique the user's data is in encrypted form in cloud and evaluating is done on encrypted data, the privacy manager make readable data from result of evaluation manager to get the correct result. In obfuscation data is not present on Service provider's machine so there is no risk with data, so data is safe on cloud, But this solution is not suitable for all cloud application, when input data is large this method can still require a large amount of memory[2]. In [3], the authors present procedural and technical solution both are producing solution to accountability to solving security risk in cloud in this mechanism these policies are decided by the parties that use, store or share that data irrespective of the jurisdiction in which information is processed. But it has limitation that data processed on SP is in unencrypted at the point of processing so there is a risk of data leakage. In [4], the author gives a language which permits to serve data with policies by agent; agent should prove their action and authorization to use particular data. In this logic data owner attach Policies with data, which contain a description of which actions are allowed with which data, but there is the problem of Continuous auditing of agent, but they provide solution that incorrect behavior. Should monitor and agent should give justification for their action, after that authority will check the justification. In [5], authors gives a three layer architecture which protect information leakage from cloud, it provides three layer to protect data, in first layer the service provider should not view confidential data in second layer service provider should not do the indexing of data, in third layer user specify use of his data and indexing in policies, so policies always travel with data. In [6], authors present accountability in federated system to achieve trust management. The trust towards use of resources is accomplished through accountability so to resolve problem for trust management in federated system they have given three layers architecture, in first layer is authentication and authorization in this authentication does using public key cryptography. Second layer is accountability which perform monitoring and logging. The third layer is anomaly detection which detects misuse of resources. This mechanism requires third party services to observe network resources.

3. PROPOSED WORK

Cloud computing is a large infrastructure which provide many services to user without installation of resources on their own machine. This is the pay as you use model. Examples of the cloud services are Yahoo email, Google, Gmail and Hotmail. There are many users, businesses, government uses cloud, so data usage in cloud is large. So data maintenance in cloud is complex. Many Artists wants to do business of their art using cloud. For example one of the artist want to sell his painting using cloud then he want that his paintings must be safe on cloud no one can misuse his paintings.

There is need to provide technique which will audit data in cloud. On the basis of accountability, we proposed one mechanism which keeps use of data transparent means data owner should get information about usage of his data. This mechanism support accountability in distributed environment Data owner should not bother about his data, he may know his data is handled according to service level agreement and his data is safe on cloud. Data owner will decide the access rules and policies and user will handle data using this rule and logs of each data access have been created. In this mechanism there are two main components i.e. logger and log harmonizer.

The logger is with the data owner's data, it provides logging access to data and encrypts log record by using public key which is given by data owner and send it to log harmonizer. The log harmonizer is performing the monitoring and rectifying, it generates the master key it holds decryption key decrypting the logs, and at the client side decryption it sends key to client. In this mechanism data owner will create private key and public key, using generated key owner will create logger which is a JAR file (JAVA Archives), it includes his policies like access policies and logging policies with data send to cloud service provider.

Authentication of cloud service provider has been done using open SSL based certificates after authentication of cloud service provider user can be able to access data in JAR, log of each data usage has been created and encrypted using public key and it automatically send to log harmonizer for integrity log records are signed by entity which is using the data and log records are decrypted and access by owner. In push mode logs are automatically send to data owner and in pull mode owner can demand logs, so he can see access of his data at anytime, anywhere and he can do monitoring of his data [1].

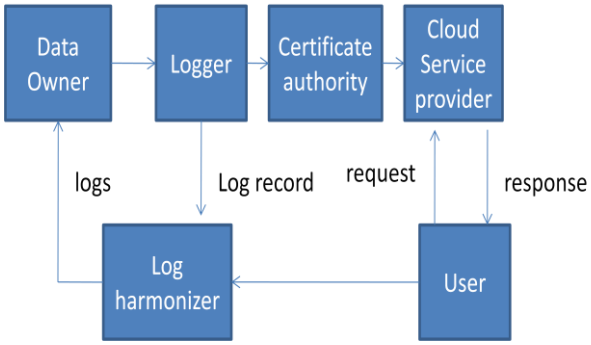

Fig 2: Accountability Mechanism in cloud

In Fig 2 working of accountability mechanism in cloud is given in this when user will access data then log of each access is created by logger and periodically sent to log harmonizer, log harmonizer send this logs to data owner and data owner can see logs and take appropriate action if he wants. State transition diagram is machine which shows no of states, machine take input from outside world and each input can produce machine to go next step. Following transition diagram shows the different states of accountability mechanism in cloud i.e. how it changes from one state to next state.

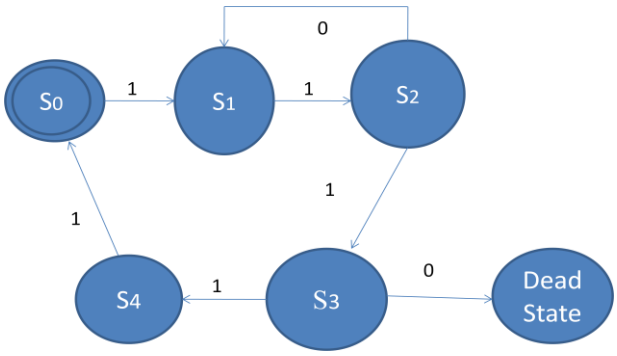

Fig 3: State Transition Diagram

Where,
0 : Unsuccessful
1 : Successful
Transition are :
S0 : Data Owner will send data to logger.
S1: Data Owner will create logger which is a jar file to store data and policies .
S2 : Authentication of CSP to JAR file.
S3 : Authentication of user.
S4 : owner can see merge log
Input: = {0, 1}
Representation of
$A = (\{S0, S_1, S_2, S_3, S4, \} \{0, 1\}, \delta, S_0, S_4)$
Input given 11011011
Expected output
δ(S0,1) = S1
δ(S1,1) = S2
δ(S2,1) = S3
δ(S3,1) = S4
δ(S4,1) = S0

In accountability mechanisms the log records are generated as access of data in jar happened then it create log record log rec (Lr).

Lr = r1, r2, r3, r4... rk.

Parameters uses for log record are
rk = (id, action, T, loc, h((id, action, T, loc)ri-1…r1), sig)

Where,
rk = log record
id = user identification
action = perform on user's data
T = Time at location loc
loc = Location
h((id, action, T, loc)ri-1…r1) = checksum component
sig = Signature of record by server

Checksum of each record is calculated and it is stored with data. Checksum is computed using hash function

$$H[i] = f(H[i-1], m[i]),$$

Where,
Compression function is $f = \{0, 1\}^n \times \{0, 1\}^b \rightarrow \{0,1\}^n$
H[i] = hash value of i^{th} log record [10], [11].

4. CONCLUSION

This paper presents effective mechanism, which performs automatic authentication of users and create log records of each data access by the user. Data owner can audit his content on cloud, and he can get the confirmation that his data is safe on the cloud. Data owner also able to know the duplication of data made without his knowledge. Data owner should not worry about his data on cloud using this mechanism and data usage is transparent, using this mechanism.

In future we would like to develop a cloud, on which we will install JRE and JVM, to do the authentication of JAR. Try to improve security of store data and to reduce log record generation time.

5. REFERENCES

[1] Smitha Sundareswaran, Anna C. Squicciarini and Dan Lin, "Ensuring Distributed Accountability for Data Sharing in the Cloud,", IEEE Transaction on dependable a secure computing, VOL. 9, NO. 4, pg 556-568, 2012.

[2] S. Pearson , Y. Shen, and M. Mowbray," A privacy Manager for Cloud Computing," Proc. Int'l Conf. Cloud Computing (cloudcom), pp.90-106,2009.

[3] S. Pearson and A. Charlesworth, "Accountability as a Way Forward for Privacy Protection in the Cloud, " *Proc First Int'l conf. Cloud* Computing, 2009.

[4] R. Corin, S. Etalle, J.I. den Hartog, G. Lenzini, and I. Staicu, "A Logic for Auditing Accountability in Decentralized Systems," Proc. IFIP TC1 WG1.7 Workshop Formal Aspects in Security and Trust, pp. 187-201, 2005.

[5] A. Squicciarini , S. Sundareswaran and D. Lin, " Preventing Information Leakage from Indexing in the Cloud," *Proc. IEEE Int'l* Conf. Cloud Computing, 2010.

[6] B. Chun and A. C. Bavier ,"Decentralized Trust Management and Accountability in Federated System," *Proc. Ann. Hawaii Int'l Conf. System Science (HICSS), 2004.*

[7] B. Crispo and G. Ruffo, "Reasoning about Accountability within Delegation," Proc. Third Int'l Conf. Information and Comm. Security (ICICS), pp. 251-260, 2001.

[8] S. Sundareswaran, A. Squicciarini, D. Lin, and S. Huang, "Promoting Distributed Accountability in the Cloud," *Proc. IEEE* Int'l Conf. Cloud Computing, 2011.

[9] D.J. Weitzner, H. Abelson, T. Berners-Lee, J. Feigenbaum, J. Hendler, and G.J. Sussman, "Information Accountability," Comm. ACM, vol. 51, no. 6, pp. 82-87, 2008.

[10] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C. John Wiley & Sons, 1993.

[11] Praveen Gauravaram, John Kelesy, Lars Knudsen, and Soren Thomsen, "On Hash function using Checksums".

[12] Ryan K L Ko, Peter Jagadpramana, Miranda Mowbray, Siani Pearson, Markus Kirchberg, Qianhui, "TrustCloud: A Framework for Accountability and Trust in Cloud Computing" HP Laboratories, pp 1 – 7, HPL-2011-38