DATA SHARING: EFFICIENT DISTRIBUTED ACCOUNTABILITY IN CLOUD

Anuradha. P and D. Kamalin

Dept. Of Information Technology, Madha Engineering college, Kundrathur- 600069, India

Abstract:

Cloud computing involves large number of computers connected through a communication network such as the Internet. Now a day's cloud computing services are emerging as a major trend in data storing and distributing data to geographically disconnected users. Now most of the persons are using Cloud computing technology. The data processed on clouds are often outsourced, leading to a number of issues related to accountability, including the handling of personally identifiable information. To solve this problem a novel Cloud information accountability approach is proposed here. The data owners enclose their encrypted data as a JAR in the outer JAR that contains the access policies. The access policies are written by owner for various users privileges. Whenever a user requests the data from the cloud, his entity details are verified by the outer JAR and the service is given to the user according his privilege. Whenever an access is made, the log harmonizer in the JAR creates a log file about the usage of data. These log files are stored in the JAR itself. These log files can be sent to owners in two modes: 1) Pull mode and 2) Push mode. In pull mode, the data owner manually retrieves the log files from the JAR. In push mode, the log harmonizer sends the log files to the users in a regular interval or if any attack was made. Third party auditor is introduced between data owner and cloud service provider which reduce the burden of data owner to audit the data in the cloud and it also make the data owner free from worrying about the data lose in cloud storage. Thus the proposed system allows the data owner to not only audit his content but also enforce strong back-end protection if needed.

Index Terms –

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

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 anywhere 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]. Cloud computing is the delivery of computing and storage capacity as a service to a heterogeneous community of end-recipients. The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure. Cloud computing entrusts services with a user's data, software and computation over a network. There are three types of cloud computing:

- Infrastructure as a Service (IaaS),
- Platform as a Service (PaaS), and
- Software as a Service (SaaS).

Using Infrastructure as a Service, users rent use of servers (as many as needed during the rental period) provided by one or more cloud providers. Using Platform as a Service, users rent use of servers and the system software to use in them. Using Software as a Service, users also rent application software and databases. The cloud providers manage the infrastructure and platforms on which the applications run. While enjoying the convenience brought by this new technology, users also start worrying about losing control of their own data. Such fears are becoming a barrier to cloud services provider [8]. Two issues are there

- Data handling can be done by the cloud service provider (CSP) to other cloud users (i.e. entities) and these cloud users may hand over to the other users so on.
- Cloud made the users whenever they can join and leave the cloud in the flexible manner.

These issues make the data handling as more complex and tedious task in the cloud. To overcome these problems introduced a Cloud Information Accountability (CIA) framework, based on the notion of information accountability. Third Party Auditor is used to reduce the burden for data owner to audit the data. IBE technique is used for encrypting purpose. Third party auditor is placed in between the data owner and the cloud service provider. Under the Database as a service, this is having four parts which are

- 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 there, 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]. In the general system model explained by Zhiguo Wan in paper [1], the cloud computing system consists of five types of parties: trusted authority, domain authorities, data owners, data consumers, and cloud service provider. The cloud service provider manages a cloud to provide data storage service. Data owners encrypt their data files and store them in the cloud for sharing with data consumers. To access the shared data files, data consumers download encrypted data files of their interest from the cloud and then decrypt them. Each data owner/consumer is administrated by a domain authority. A domain authority is managed by its parent domain authority or the trusted authority. Data owners, data consumers, domain authorities, and the trusted authority are organized in a hierarchical manner [1]. In this first the data owner will set the policies for the data which he/she wants to place in cloud and send it to cloud service provider (CSP) enclosed in JAR files, any access to the data will be automatically checked for its authentication and logs record for each data item will be created and sent to data owner for monitoring the data usage.

RELATED WORKS

In this section review related works addressing security in cloud. Security issue are 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.

PROPOSED METHOD

Cloud computing is a large infrastructure which provide many services to user without installation of resources on their own machine. 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, proposed one mechanism which keeps use of data transparent means data owner should get information about usage of his data.

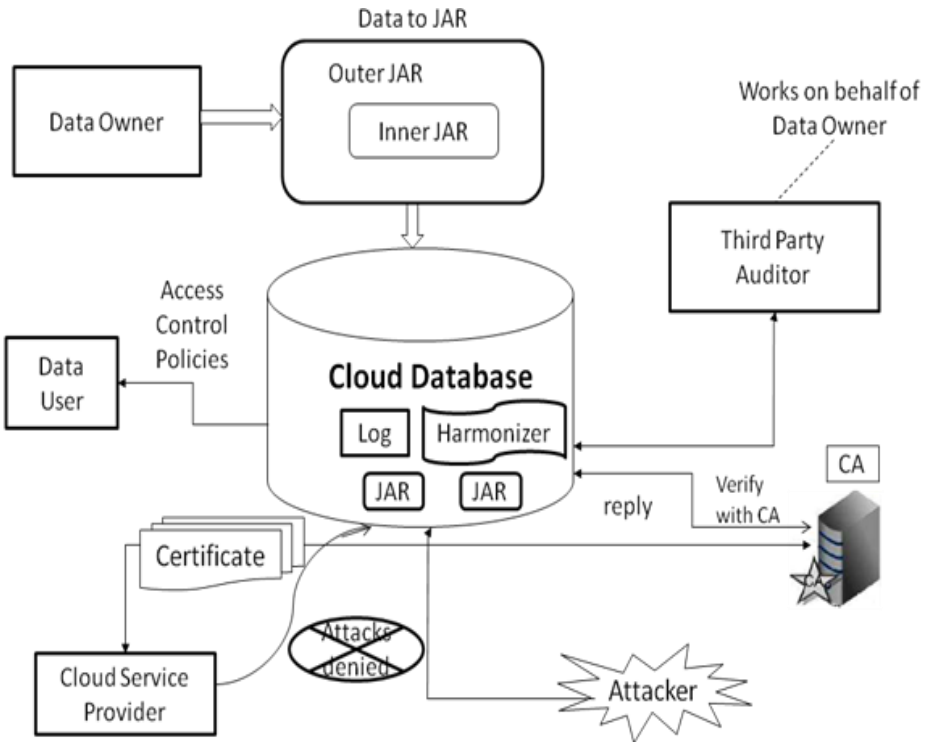

Fig.1. Overview of proposed system

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. The proposed system methodology is shown in Figure1 below,

At the beginning, each data owner creates a pair of public and private keys. The data owner encrypts his/her data using the generated key and stores it in Inner JAR. Then he writes access control policies for various users and stores them in an Outer JAR. The inner JAR is protected by Outer JAR. Once the JAR is ready he stores it in cloud server. Here a third party auditor is introduced to track the usage of data stored in cloud and log records. So the third party auditor works on behalf of data owner and monitors the log history. Thus the overhead of data owner is reduced here. The cloud users access the data items stored in JAR according to the access policies they are able to meet [11].

For example If the data owner specify a policy such that the user from London can only view my data, then obviously the user has only view access and he can't download the data. Likewise several access policies are written for users. If the cloud service provider himself wants to access the data items, then he needs to get a certificate from certificate authority and then has to furnish it to the JAR stored in cloud. Here the attacker creates three types of attacks. Man in middle attack, Disassembling JAR attack, Copying Attacks.

In Man in middle attack, the attacker steals the identity certificate in middle way and uses the same to masquerade as a legal CSP. But our proposed system prevents this attack with the help of timestamp present in the certificate.

In Disassembling JAR attack, the attacker corrupts the log files by entering fake details in the log files. But our proposed system prevents this attack with the help of log harmonizer and fixes the corrupted log. In copying attacks, the attacker creates more copies of Jars. This attack can't be blocked but the information about this attack is sent to third party auditor. The fact is whenever an attack is made, the log about the corresponding attack is sent to the third party auditor by the log harmonizer of the corresponding JAR.

Modules

The major buildings modules of proposed systems are Four. They are.

1) Cloud Entities and JAR files construction
2) JAR files storage and user access
3) Third Party Auditing and Handling Man-in-middle attacks
4) Handling Disassembling JAR and copying attacks

CLOUD ENTITIES AND JAR FILES CONSTRUCTION:

Initially GUIs for data owner, cloud service provider, third party auditor, certificate authority GUI and user accounts are created. The users construct data files and access polices for users. Then he generates encryption keys and encrypts the data. The encrypted data and access policies are sent to cloud server in the form of JAR. Now the jar is stored in any of the cloud servers specified in the access policy. The access policies differ for users from different geographical location. The data owner uploads their data in the cloud server. The new users can register with the service provider and create a new account and so they can securely upload the files and store it. For the security purpose the data owner encrypts the data file and then store in the cloud. The Data owner can have capable of manipulating the encrypted data file. And the data owner can set the access privilege to the encrypted data file. To allay users' concerns, it is essential to provide an effective mechanism for users to monitor the usage of their data in the cloud. For example, users need to be able to ensure that their data are handled according to the service level agreements made at the time they sign on for services in the cloud. The cloud service provider manages a cloud to provide data storage service.

Data owners encrypt their data files and store them in the cloud with the jar file created for each file for sharing with data consumers. To access the shared data files, data consumers download encrypted data files of their interest from the cloud and then decrypt them. Creating the jar file for every file upload, the user should have the same jar file to download the file. This way the data is going to be secured. The logging should be decentralized in order to adapt to the dynamic nature of the cloud. More specifically, log files should be tightly bounded with the corresponding data being controlled, and require minimal infrastructural support from any server. Every access to the user's data should be correctly and automatically logged. This requires integrated techniques to authenticate the entity who accesses the data, verify, and record the actual operations on the data as well as the time that the data have been accessed. Log files should be reliable and tamper proof to avoid illegal insertion, deletion, and modification by malicious parties. Recovery mechanisms are also desirable to restore damaged log files caused by technical problems. The proposed technique should not intrusively monitor data recipients' systems, nor it should introduce heavy communication and computation overhead, which otherwise will hinder its feasibility and adoption in practice.

JAR FILES STORAGE AND USER ACCESS:

The cloud server stores the data in their server. The data owner meanwhile delegates the data to third party auditor for auditing purpose. Now the users login to the cloud application. Then they enter query strings to access data. After that the users can download or view the data according to the specified access policies.

THIRD PARTY AUDITING AND HANDLING MAN- IN-MIDDLE ATTACKS:

The log harmonizer generates log files whenever user access data from the cloud. The log files are periodically sent to the third party auditor by the log harmonizer. The log file contains time of access, file access status IP address of user machine. The third party auditor can also access the log files from the harmonizer by running pull mode. The cloud service provider requires certificate from the certificate authority to access the jar from the cloud. The certificate contains a timestamp for duration of data access. The cloud service provider then submits the certificate to jar for verification. At that time, the attacker gets the copy of the certificate. The jar verifies the certificate and allows the service provider to access the data. After the session of service provider ends, the attacker submits the same certificate to jar. Now the jar denies the access since the timestamp is invalid.

HANDLING DISASSEMBLING JAR AND COPYING ATTACKS:

The attacker disassembles JAR and corrupts the information present in it. If the log file is edited, then the harmonizer fixes the changes made in the log file. If the attackers create copies of JAR, then notification about the attack is sent to third party auditor. Whenever an attack is detected, the log harmonizer sends the notification information about the attack to third party auditor.

ADVANTAGES

1) One of the main innovative features is that it is a light weight architecture, only data can be stored are given by the actual files and associated logs.
2) JAR act as a compressor of the files and multiple files handled by the same logger component, used to handle more than one file, results of storage transparency.
3) It allows the data owner to not only audit his content but also enforce strong back-end protection if needed.

TOOLS USED FOR IMPLEMENTING CLOUD

In the proposed model we are using the following tools:

- **EUCALYPTUS CLOUD**

The Eucalyptus Cloud platform is open source software for building AWS-compatible private and hybrid clouds. Eucalyptus supports Amazon web services EC2 and S3 interfaces. It pools together existing virtualized infrastructure to create cloud resources for compute, network and storage.

- **AMAZON EC2**

Amazon Elastic compute cloud (EC2) is a central part of Amazon.com's cloud computing platform, Amazon web Services (AWS). EC2 allows users to rent virtual computers on which to run their own Computer Applications. They are designed for control and management of VM instances, EBS volumes, elastic IPs, and security groups and should work well with EC2 and Eucalyptus [10].

CONCLUSION

Presents an 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 data on cloud using this mechanism and data usage is transparent, using this mechanism.

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. Feigen-baum, J. Hendler, and G.J. Sussman, "Information Accountability," Comm. ACM, vol. 51, no. 6, pp. 82-87, 2008.

[10] EucalyptusSystems, http://www.eucalyptus.com/,2012.

[11] 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.