International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
International Conference and Workshop on Communication, Computing and Virtualization (ICWCCV 2015) – www.ijais.org

An Approach to Ensuring Distributed Accountability in Cloud Computing – A Survey

Shubhangi P. keni
M.E. Student, Dept. of Computer Engg,
Shah and Anchor Kutchhi Engg. College,
Chembur, Mumbai-400088, India

Shahzia Sayyad
Professor, Dept . of Computer Engg,
Shah and Anchor Kutchhi Engg. College,
Chembur, Mumbai-400088, India

ABSTRACT
Cloud computing is the emerging technology which makes it possible for you to access your data from anywhere at any time. Where in traditional method computers setup requires you to be in the same location as your data storage service. A distinct feature of the cloud services is that user's data are remotely proceeds in unknown machines that user do not own or operate. Data handling goes through complex and dynamic hierarchical chain through various abstraction layers. While enjoying this new emerging technology, user's fear of losing control of their data particularly (health & financial data, etc.).Hence cloud computing can become a significant barrier to the wide acceptance of cloud services. To solve this problem data should be shared and used in cloud environment with the help of trust components such as privacy, security, accountability & auditability techniques in order to achieve the customer's confidence level and long term customer relationship. This paper discusses the cloud accountability life cycle through different phases of the log data. And various frameworks has been addressed to the automated logging for ensuring accountability and distributed auditability performed by any entity, carried out at any point of time at any cloud service provider**.**

Keywords:
Accountability, auditability, cloud computing, data sharing, security, privacy, trust.

1. INTRODUCTION
Cloud computing is the utilization of hardware& software resources like (RAM, CPU cycle, storage, network bandwidth, virtual machine, web server & database etc.), which are used for computation delivered as a service over a network or Internet with minimum cost effect and IT services. While enjoying the convenience & uniqueness of cloud computing such as on demand self-service, rapid elasticity, pay as you need, broad network access& multi-tenant nature of working environment etc., it is very much difficult to keep control on data and process which are carried out remotely on someone's physical or virtual machine's location that user don't know. The data processed on clouds are often subcontracted, leading to a number of issues related to accountability including the handling of personally identifiable information. People find the loss of control over data Or information which leads to lack of trust in cloud. It is essential to provide an effective mechanism for users or authorized third party to monitor the usage of their data in order to achieve trust in the cloud. There are four trust components such as privacy, security, accountability and auditability in cloud computing [1].

Preventive controls to solve privacy & security problem has been focused or researched by using many encryptions and hashing techniques. But detective controls to solve accountability & auditability problem is still little focused or researched. Accountability is the verifying or checking the authorization polices(e.g. user's roll based read, write, copy, & execute permission over the data or process), which will help to built-up the trust relationship with the customer in order to enhance the business activity in the cloud computing environment.

The end user is allowed to access the data as per their access privileges, which they specifies while registering to access the data in the cloud and authentication is provided and data usage will be verified. Auditing is the checking of the log files to generate summarized report for understanding the security breaches which will help to take major control on the usage of data. Auditing plays an important role for monitoring the irregularities in cloud environment. It can be carried out periodically or randomly or as per the requirement of the stakeholders or cloud Server Provider. Accountability & auditabilityboth plays an important role of keeping an eye on ongoing process or activity in real time carried out in cloud computing in order to empower the trust relationship between the data owner, end user and cloud service provider. There are many situations where accountability is needed,

for example any professional photographer or artist plan to sell his/her photographs or art by using the Sky-high cloud services for their business in the cloud, where they may need following requirements such as[2].

- To verify whether the end user & Cloud Service Provider is authorized or not.
- To verify whether user is the paid or potential buyer.
- To ensure the service level agreement is honored or not [3].
- To verify the geographical access to shared data or information is there or not [2].
- To ensure the service provider do not share data with other service provider.
- In case any dispute arises with client, you may need all the relevant information about the client to solve the problem.
- Finding answerable authority for any fault or incident that occurs.

- finding the attack which destroys the customer's software to steal valuable data or to launch the spamming mail attacks or DOS attacks etc. [3].
- Finding threats to cloud computing such as malicious insiders, insecure application programming interfaces, data loss or leakage, abuse use of cloud computing, shared technology vulnerabilities, account, service and traffic hijacking [1].

The rest of the paper is organized as follow, section2 discusses cloud accountability life cycle. Section3 presents Trust Cloud framework [1] and cloud information accountability framework [2]. Section 4 describes conclusion and section 5 describes the references.

2. THE CLOUD ACCOUNTABILITY LIFE CYCLE [1,4].

Accountability is the verifying or checking the authorization polices. There are seven phases of Cloud Accountability Life Cycle given below.

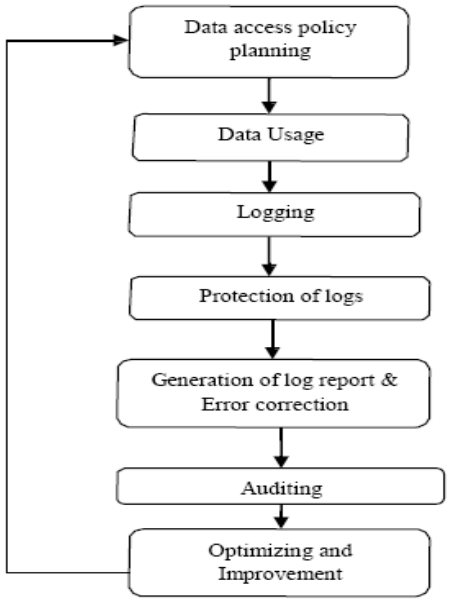

Figure 2.1 Cloud Accountability Life Cycle [1, 4]

A) Data Access policy planning

Data owner or cloud service provider will set the data access policies to decide what information to log and which events to log [1, 2].Generally four important groups of data must be logged.

- Event Data–a sequence of actions and relevant information.
- Performer Data- the data user or computer Components (e.g. worm) which trigger the incident.
- Timestamp Data- the time and date incident took place or location.
- Location Data-both virtual & physical (network, memory& storage etc.) server addresses at whichevents took place or the location.

B) Data usage

Any misuse activity or irregularities carried out should be sensed and triggered to logging in CSP's, while data Usage carried out (in real time).Accountability tools need to be able to trace from lowest-level system read/write calls all the waysto the irregularities of high-level workflows hosted in virtual machines in different physical server and location also there is need to trace the network packets routes within cloud.

C) Logging

There are two types of logging. One is file-centric point of view logging through system layer and another is data-centric point of view logging though data layer. File centric point of view logging is performed on both virtual and physical layer in cloud. Life span of logs needs to be considered within cloud, the details of data usage to be logged & location of storage of the logs.

D) Protection of Logs

After logging is done, there is need to safe keeping of log files i.e. protecting the integrity of logs from unauthorized access and ensuring that they are corrupt- free. Encryption technique is needed to protect the logs. And proper Backup mechanisms also needed to prevent the loss of logs or corruption of logs. This phase plays important role for the next reporting & error correction phase.

E) Generation of log report & error correction

Data will be used by user and logs of each record will be created to generate the reports. Suspected irregularities are also flagged to the end user. In case any log corruption is occurs then error correction is done by replay, backup, rollback, recovery and restoring the data. Reports may cover a large scope, scale and size of logging. Log recording or monitoring can be from virtual & physical server histories within the cloud. It can be from OS-level read/write operations of sensitive data, or high level workflow i.e. from user and user management audit trails.

F) Auditing

Logs and reports are checked and potential irregularities highlighted. The checking can be performed by auditors or stakeholder or trusted third party. Automated auditing process will be permitting auditors to detect irregularities very efficiently.

G) Optimizing & improvement

Suspected area and security loopholes in the clouds are removed or improved. This phase will help to take all corrective controls and governance of the cloud processes. This phase plays important role to define certain rule and regulations as well as data access policies [1, 2, 4].

3. FRAMEWORKS / SYSTEM ARCHITECTURE

Trust Cloud framework [1] and Cloud Information Accountability framework [2] has been studied to understand the distributed accountability in cloud computing environment. These frameworks are used to keep track of the actual usage of the user's data in the cloud.

3.1 TrustCloud Framework

Trust Cloud framework addresses accountability in the cloud. Trust Cloud framework consists of the abstraction layers of accountability such as system layer, data layer, workflow layer laws & regulation layer & policies layer.

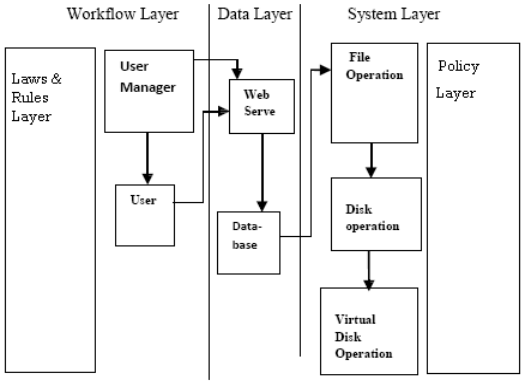

Figure 3.1 Abstraction layer of Accountability in cloud Computing [1].

Figure 3.1 shows the abstraction layers for which types of logs can be created for the accountability purpose. Since excess amount of logs types range from a system level to a workflow level. Log scale size and scope need to be decided by the auditor. This framework is allowed to identify accountability needed at all the level of granularity. Multi-layered logging focuses on file & data centric rather than system centric

A) System Layer

System layer performs file-centric logging within operating system, file system & cloud internal network [1,5].Operating system and event logs are the most common types of logs associated with the cloud computing at the moment. File system is technically part of the operating system. From the file system read/write calls log files can be extracted from virtual and physical memory location. It is providing more information for tracking the event or incident. Cloud internal network includes many physical and virtual computing resources such as servers, memory & Storages etc. which are spread over the different location. It is also needed to monitor the network logs. Logging information over the operating system, file system and the cloud internal network not only used to study the tracing as well as to understand the file's life cycle (i.e. .creation, modification, duplication, retrieval & destruction etc.).

B) Data layer

Data layer performs Data-centric: logging where data is truly belongs to the user within provenance logger & consistency logger. Provenance logger enables reasoning about the origin, collection or creation, evaluation and use of data which will help to trace the history of data i.e. its origin. Provenance logging must fulfill the following criteria such as to ensure that the logs themselves are tamper-free, logs should be eventually consistent and complete followed by ACID properties of database, to be scalable, to be persistent over the long term, to be allow access based on roll with different access privileges, to be

efficiently accessible & to be transparent. Consistency logger must eventually consistent to allow for rollback, recovery, replay, backup, & restoring the data.

C) Workflow layer

The Workflow layer focuses on the audit trails and the audit related data found in the software services in the cloud. The workflow layer aims to ensure proper governance of cloud application, empower continuous auditing and manage proper governance of cloud application, empower continuous auditing and manage the accountability services composed as business processes via automated auditing, patch management auditing, accountability of services. To achieve the accountability of services logging should be take care of the concerns on a component such as input or Pre-processing, processing and post processing.

D) Policy, laws and rules layer

Policy, laws and regulation require information to logged on what, when, why, where, how, and by whom data processing takes place. What refer to data classification with different policies and legal rule. Data classes to consider might include Non-Personally Identifiable Information data, anonym zed data, Personally Identifiable Information, sensitive Personally Identifiable Information, and payment card industry regulated data. Why refers to purpose of data processing action and for which the processing of given Personally Identifiable Information (PII) data item is permitted may need to be recorded. When refers to timestamps. Timing information is also needed for compliance to laws & policies concerned with the data preservation. It may reduce the information that need to be recorded. Where refers to the geographical location matters to legal point of view. It can be ascertain in cloud where data is, there can be multiple copied of data, so the physical location of storage and the occurrence of cross border data transfer may need to be recorded. How refers to some laws & policies restrict how data is handled. For e.g.Processing of Payment Cards Industry regulated data may require encryption & other safeguards. Who refers to policies restricts data access to particular set of authorized user, individual or by role based access. There may need to record the corporate identity of partners or cloud service providers to which data is transmitted which will be help to assist action required by policies if provider goes out of business or has a data breach.[1]

3.2 Cloud Information Accountability (CIA) Framework [2, 7].

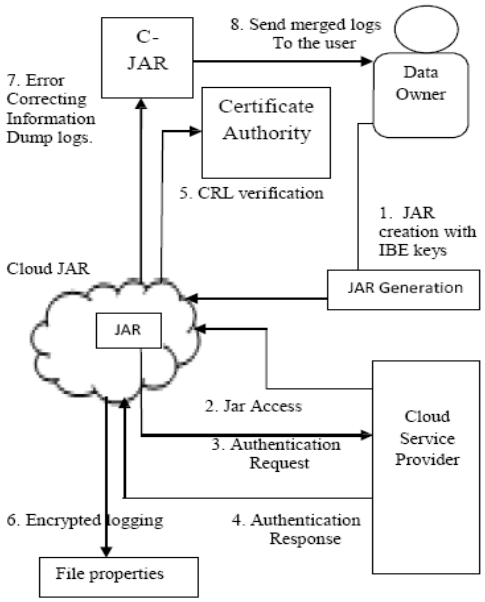

Figure 3.2 show the cloud information accountability framework [2, 6, 7].

Figure 3.2 shows the Cloud Information Accountability framework used to conducts automated logging and distributed auditing of relevant access performed by any entity, carried out at any time at any clouds service provider. There are two major components of the Cloud Information accountability such as logger & log harmonizer. The logger is strongly coupled with user's data (either single or multiple data items). The logger can be configured to ensure that data access & usage control policies associated with data are honored or not. Its main tasks include automatically logging access todata items that it contains, encrypting the log record using the public key of the content owner, and periodically sending them to the log harmonizer. Log harmonizer is responsible for auditing. Feature of CIA Framework is the logger & log harmonizer component both are light weight & portable JAR files which provide automated logging [2, 6, 7]. The CIA framework is explained by the step given below.

Step 1 – Data set with Access policies.

- At the beginning, each user creates a pair of public and private keys based on Identity based encryption. (IBE).
- Using the generated key, the user will create a logger component which is a JAR file, to store its data items.
- The JAR file includes a set of simple access control rules specifying whether and how the cloud servers and Possibly other data owner or stakeholders are authorized to access the content itself.

Step 2- Jar Access

- after step 1, data owner sends the JAR file to the cloud service provider that he subscribes to.

Step 3-5 Certificate Authority verification

- To authenticate the Cloud Service Provider to the JAR, OpenSSL based certificates is used, wherein a trusted certificate authority certifies the CSP.
- If end user requested for the data access, then authentication is done by using SAML-based certificates, wherein a trusted certificate authority certifies the end user.
- Once the authentication succeeds, the service provider or end user will be allowed to access the data as per the enclosed policies.

Step 6- Protection of log records

- Whenever there is access to data, the JAR will automatically create log record, encrypt it using public key distributed by data owner, and store it along with data so that logs will be tamper free.

Step 7- Error correction

- In addition, some error correction information will be sent to the log harmonizer to handle possible log file corruption.

Step 8- Auditing

- The encrypted log files can later be decrypted and their integrity can be verified.
- These logs can be accessed by the data owner or stakeholder at any time for auditing purposes with the aid of the log harmonizer.
- There are two auditing strategies such as pull and push mode.
- The Push mode:

  The push mode refers to logs being periodically sent to the data owner or trusted third party.

- The Pull mode:

  The Pull mode refers to an alternative approach of auditing whereby the data owner or trusted third party can retrieve the logs as per their need.

4. CONCLUSION

Cloud makes IT resources as a service. But security breached is key barrier to wide acceptance of cloud computing. To solve such trust issues some preventive controls as well as detective controls are need to be focused or researched. Preventive controls are used to lessen the occurrences of action from continuing further or taking place at all like access policy that issued to govern the user to modify or read a file or database, or network and host firewall used to blocks all unauthorized sites but allowable activity etc. Detective controls are used to monitor or record to identify the ongoing incident against the privacy or security policies & procedures like an intrusion detection system on a host or server or network or security audit trails, logs etc. Accountability and auditability strategies make cloud more trusted by lessen the security loopholes.

5. REFERENCES

[1] "Trust cloud: A framework for accountability and Trust in Cloud Computing. Ryan K L Ko, Peter Jagadpramana, Mirandra Mowbray, Siani Pearson, Marhus Kirchberg, Qianhui Liang, Bu sung Lee, HP Laboratories, HPL-2011- 38

[2] "Ensuring distributed Accountability for Data Sharing in the Cloud. " Smitha Sundareswaran, Anna C.Squicciarini , member, IEEE, and Dan Lin, IEEE transaction on dependable and secure computing Vol. 9 No.4 Year 2012

[3] "A case for Accountable Cloud" Max Planck Institute for software System (MPI-SWS)

[4] "Distributed Accountability for Data Sharing in Cloud" Snehal Suryawanshi, Anant M. Bagade, International Journal of computer Applications (0975-8887) Volume 59- No,8, December 2012

[5] "Flogger: A File-Centric Logger for Monitoring File Access and Transfers within Cloud computing environment", K L Ryan Ko, Peter Jagadpramana, Bu Sung Lee, HP Laboratories,HPL-2011-119

[6] "Review on Techniques to Ensure Distributed Accountability for Data Sharing in the Cloud" H.Arun, R.Nilam, S.Purva, International Journal of Advanced Research in Computer and Communication Engineering, volume 2.,October 2013,

[7] "Ensuring Distributed Accountability for Data Sharing in the Cloud" Pankaj Kumar Singh, Ajit Kumar *&R.Krthikeyan, International Journal of Advanced Researchin Computer Science and Software Engineering, volume 3, March 2013,www.ijarcsse.com.*