Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 50 (2015) 543 – 550

Procedia Computer Science

2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

Towards an analysis of data accountability and auditing for secure cloud data storage

Prassanna.J, Punitha.K, Neelanarayanan.V

School of Computing Science & Engineering,
VIT University, Chennai, 600 127, India

Abstract

Cloud technology, runs off-premise, provides an extensive information technology support and service that can offer as an on-premise version if on demanded as a virtualized solution for customers' technological requirements. Cloud computing put together the comprehensive solutions by integrating the different technology that provision Software, Platform and Infrastructure as services. Cloud, an integrated technology that delivers a complete, open and flexible solution, leads to a critical concern over its trust. The extensive use of virtualization in cloud especially in data storage capability put forth many security concerns to the customer. It is the time to prove the customer about the strength of the cloud security through enabling their data accountability and auditing mechanism. In this work we going to provide a systematic analysis of the data accountability mechanism and auditing approaches existing now in cloud. To strengthen the customers trust over the cloud technology and service, we provided a clear enhanced view on the cloud data security and also analyze the secured way of utilizing the advantages of cloud computing.

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

1. Introduction

Cloud Computing provides a new standard to support the worldwide computing demand for Information Technology services. It breaks the limitations on network service provisioning as any computing requirements such as software, hardware and infrastructure will be delivered as services through the network. The standard also provides flexibility to the customer's instant, dynamic and wide array of service requirements.

In the regular or conventional computing system, the processing or the user data are stored in the users'

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)
doi:10.1016/j.procs.2015.04.101

machine and the computing capability is also up to the users' computer, but in contrast the cloud computing framework the data storage and the processing power are rendered as a large scale virtualized facility as a service by a well-organized and well managed cloud service providers. The cloud computing model provides the user easy to deploy, simple to operate and maintain their information processing system and it's in great extend reduce the costs when the system needs to increase its reliable efficiency.

As per the service extension the cloud computing is categorized as public cloud services, private cloud services, hybrid and community cloud services. The models of the cloud computing services are classified as on the basis that how different category of resources are rendered as a computing service by the cloud service provider. The model of cloud services as rendered by the cloud computing service providers are classified as Infrastructure (Iaas), Platform (Paas), Software (Saas) or Composition as a Service (CaaS).

2. Classification of Cloud Computing

Nowadays the name "Cloud "is a buzz word in modern computing environments. Even though the cloud computing is a new and fast growing technology, the fact is that early classes of this technology existed already before this. Earlier form of cloud with added technology and added constructs have been developed to make the present cloud system. Cloud version 1.0 is the distributed computing, cloud like technical invention ensued from the Transmission control protocol and Internet protocol layers, in which the networked devices exchange information between each other by adapting with the protocol spec with the remote entities even unknown about the locality of the entities and unknown about the other end of the communication. The generalization of data on the World Wide Web in the internet is that where the information are posted on the web's unknown to the user about where they are posted and also the data retrieved unknown about the location of the data can be measured the next version of cloud that is 2.0. The present version of the cloud is 3.0, which as catchy as it provide anything as a service. In addition to the software service provisioning it provides complex computing infrastructures, modern platforms, and mass storage as services through the network without complicating the user about the locality of the service provisioning. Cloud computing makes an extensive service support from individual requirements to an enterprise requirement as per the demand arise. As per the service extension the cloud computing is categorized as public cloud services, private cloud services, hybrid and community cloud services.

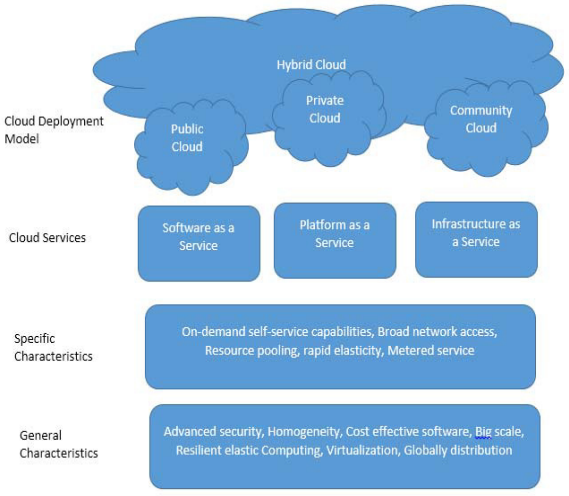

Fig. 1. The cloud computing model

- A Private cloud is a class of service that provides services to a group of limited user such as to corporate or organization requirements. It is owned by an organization or corporation. The private cloud computing services are dedicated to that organization that is to the users of that organization. The blend of Virtualization along the distributed computing provides the on demand computing capabilities as a cloud service to the organization in turn it allows the organization's network and datacenter administrators

provide effective services to the users within the organization. Any cloud model that's designed by a corporation or an organization to deliver its specific facilities are example for private cloud.

- A Public cloud is a general cloud computing model. In this model the cloud is owned by a specific service provider. The service provider provision the resources such as the utility software or the required storage spaces to the public users through the internet. The service offered by the public cloud service providers may be cost free or the services are rented for the users or customers according to their utilization. Some of the examples of public cloud include Google Drives, Apple iCloud, Amazon Elastic Compute Cloud, Windows Azure Service Platform.
- A Community cloud is a pay for share infrastructure that is partaking among various organizations from a particular group or team with common or similar computing business. The cloud services rendering by the cloud service provider shall be utilized by a group of people from different Workgroup or company those who are working on a common computing business because of to reduce the cost of the service and extracting the essence of dynamic facilities provisioning and sharing of cloud. The Lockheed Martin offers a secure community cloud the SolaS that render a highly secure, multi-tenant atmosphere for the government and governing organization seeking the benefit of cost efficient shared infrastructure over cloud is an example of a community cloud.
- The Hybrid cloud is the facility that offers the company or organization to combine its private cloud computing capability with the public cloud to utilize the computing services provided in public cloud for efficient costing or pay free without compromising any critical information or policies. The major go for the hybrid cloud is to rend additional infrastructure in the demand rising situations. The goal behind the hybrid cloud is to combine services and facilities such as data centers across different model of cloud.

The Models of the Cloud Computing Services are classified as on the basis that how different category of resources are rendered as a computing service by the cloud service provider. The model of cloud services as rendered by the cloud computing service providers are classified as Infrastructure (Iaas), Platform (Paas) , Software(Saas) or Composition as a Service (CaaS). The Fig. 2 represents the relationship between the layer of the application stack and the diverse cloud service models.

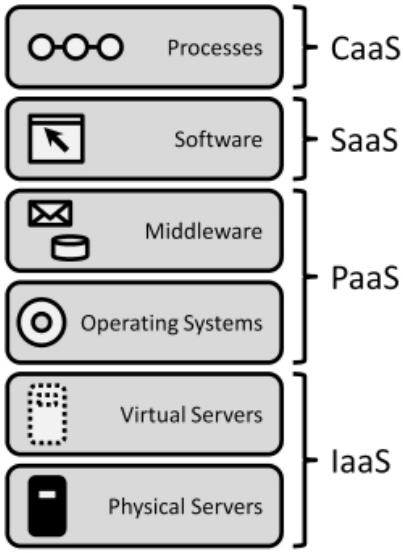

Fig.2. Relationship between Layers of the application stack and the diverse Cloud Service Models [9]

- SaaS Cloud uses the World Wide Web to deliver application services that are managed by an intermediary vendor and whose interface is accessed by the end-users side. The organization resource planning software, business processing software, accounting software, collaborative software applications like customer relationship management (CRM), enterprise resource planning (ERP), and Tally are offered by Software as a Service. Examples for Software as a cloud service are DeskAway,

Impel CRM and Wipro w-SaaS. The cloud service user does not have any control over the basic infrastructure.

- PaaS Cloud service is a model that in addition to the software the hardware resources and tools running on a cloud infrastructure rendered as a service to the end-users or developers. It provide the infrastructure for the development through PaaS to develop and modify applications. For the developers the PaaS renders the development platform, testing tools, and deployment facilities over required applications as required, as fast, as simple effortless, and as cost-effective such as avoiding the necessity to own the underlying layers of hardware and software infrastructure. Amazon web services, Appistry's CloudIQ, AppScale, openstack, Google, LongJump are some of the examples of Platform as a service Cloud**.** There it does not provide any control or domination over the underlying basic cloud infrastructure or the platforms, for the users or the customer. Platform as a Service facilitates the user to build applications using software components that are controlled by an intermediary vendor. Platform as a Service is extremely scalable, and users don't have to concern about software upgrades or having their service go down in maintenance.
- Infrastructure as a Service cloud is a standardized, provision framework in which a company or organization outsource the on-demand computing equipment required to support process such as mass storage, efficient hardware, high performance servers and network elements. The IaaS cloud service provider possesses the computing equipment and is responsible for holding, keeping, processing and maintaining the cloud infrastructure. The utilization charges as per the use basis is charged to the client. As an alternative to purchase for software, storage, servers, or network equipment, end-users can pay money for these as a fully outsourced service that is generally billed according to the amount of facilities utilized. BlueLock, Amazon web services, AT & T cloud computing services, Rackspace are some example of Infrastructure as a service cloud.

3. Security disputes in cloud computing

3.1 Straight Forward Security Challenges

The security issues in the conventional information and communication systems also employ to the cloud computing system. The practice of cloud computing gets into modern attack challenges that will cause attacks either feasible or lighter to handle it. The availability regardless of time or day of the service rendered by the cloud is also a major worry, since if the cloud service is cut off or disturbed then it risks more customers than in the conventional cloud service framework. The security of the cloud virtual machines also a major concern. The Virtual machine manager such as the hypervisor and the virtual machines employed in a cloud service provider may also undergo vulnerabilities.

3.2 Cloud Specific Security Challenges

The important characteristics of the cloud services that elaborates the relation with and differences from the conventional computing models are On-demand provisioning, it refers to the service rendered by the cloud service provider that empower the provision of cloud resources on demand by the user, Wide network access, Resource pooling, Rapid elasticity, it refers to the scalable service provisioning or the capability to provide quickly scalable services and Metered service , it is pay as per use.

4. Cloud data accountability

The absence of transparency in the cloud services questioning many users to utilize the maximum gains from the cloud technology. But at present there are service providers in the cloud they ensure the users by rendering simplified solutions for day to day use, it looks that a small effort has been provided to the questions on the things went wrong; revenue loss for the cloud service provider, cloud properties with user stored data are caused by the effectual authority and show off. At present the cloud service providers are demonstrated in a fashion which entails that the cloud users must lay complete trust over the cloud service provider, even though there is a fallible person in

the intertwine, such belief in the cloud may not every time be defendable. The cloud characteristic of availability or accessibility of the services anywhere around the world bring in the demand of following with contradicting legal issues. The cloud accountability ensures the transparency in the cloud model, the responsibility of the cloud service, confidence to the cloud users and remediation while trouble. The accountability should be embedded with the present cloud architecture as an extensive technique for the essential and efficient safety insurance to include both protective and reactive actions. Cloud accountability mechanism on the go will study their needs while severe issues arises and build the protective and countermeasures can be built with it. Incorporating the accountability mechanism in the cloud will require many cloud services to incorporate the measurable legal and technical actions to back up such an efficient and effective cloud accountability. Fig. 3 shows the framework for data leakage solution in cloud.

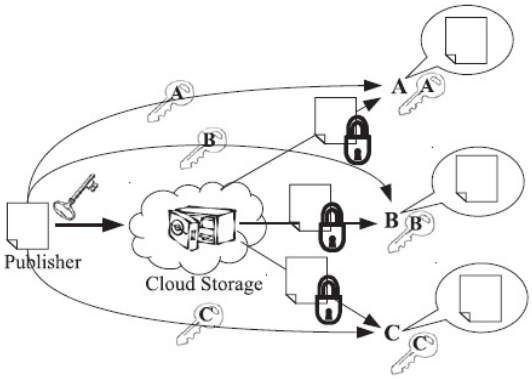

Fig. 3. A cloud data leakage prevention solution [1].

The researcher makes the accountability measures as a demonstrable mechanism implemented through the most secure cryptographic algorithms especially in the eCommerce setting [6]. A good illustrative work is presented in this field by [7]. In the work the authors admitted the policies associated with the data and proposed the executable system for accountability cloud data in a distributed setup. The proposal for designing accountability inclusive distributed computing system is given by [5].

5. Providing accountability in the cloud

Cloud accountability encourages the execution of techniques that can be implemented in practice through legal necessities and directions are converted as an efficient and effectual information safeguard mechanism. Governing laws and legal policies be inclined to concern about the information communication data level, but the process techniques for accountability can exist at different levels that admitting cloud framework and information levels. It is meaningful to design an integrated legal technique, actions, and systematic high level measures to defend this type of cloud system design. In the design of such secure cloud system it may incorporate formulated attributes to support [2]

i) Defensive accountability, using proactive controls and
ii) Counter accountability, using reactive controls.

Proactive controls can help identify whether a controlled activity extends to the system or include it. For the cloud system a proactive control is typical process that admit hazard identification and action incorporation mechanism, policy implementation (for example, promotion policies, data integrity policy, machine understandable policies, privacy or secrecy control policy, and responsibilities), trust measurement, malware mystification techniques, and characteristic directions.

Enterprises can use reactive controls to classify data privacy and data security dangers that violate the identified protocols and mechanism (for example, identifying the intruder in a computing environment, data access policy guaranteed transaction logs, language constructs, and analysing mechanisms). Cloud reactive controls include data auditing, information tracking, log recording, and supervising.

Technical procedures for implementing accountability can admit data encoding for data safety improvement,

data privacy preserving policies, and factors to help boost trust. It is also capable to depend on infrastructure and environment to preserve proper detachments, impose policies, and account data exactly. Finally, it is analysed that to attain data accountability in the cloud practicing legal guarantees through the cloud service supplying chain from cloud service providers to data accountable organizations, improved on the technical side by implementation of matching system acceptable rules transmitted with information along the cloud, incorporated hazard evaluation, confidence assurance, and inspecting audit. As a result, the data accountable mechanisms can guarantee that all the cloud users who operate data discover their responsibilities to safeguard it, regardless of the locality of the operation occurs

6. Distributed data accountability in the cloud

In the Cloud the properties of the end-users such as the personal data, the business data are kept in the remote location unknown to the user. The data processing such as transactions, accessing, updating, modifying, etc., is also processed remotely in the cloud environment. This leads to a fear to the end-user about that they can't have a hold on their data kept in the cloud. Even though cloud provides the virtual environment that makes the end-user feel comfortable in working with cloud as working with their personal environment. This becomes the major barrier to the development of the cloud technology. To build the trust on a cloud the distributed data accountability embedded cloud framework is to hold the control on the user data stored in the cloud by keep follow the existent usage resulted replication of the end-user data in the cloud system through an object focused mechanism effectively implement a logging mechanism which enveloped unitedly with the end-user data and the associated cloud data policies in the cloud [8].

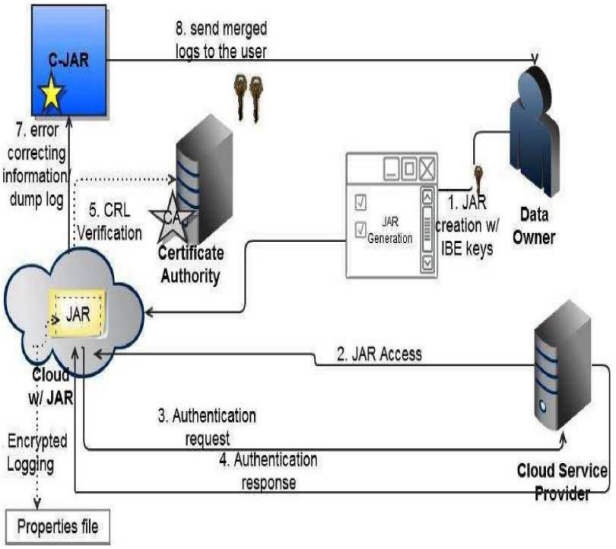

Fig. 4. Overview of the cloud information accountability framework [2]

The Cloud data accountability framework is designed to have a machine controlled logging mechanism associated with globally distributed data auditing of appropriate data access executed by whatever activity, accomplished at any point of reference at any cloud service provider. Logger and log harmonizer [8] are the components designed for the implementation. It utilizes the essence of object oriented techniques in cloud information processing. The first component logger [8] is tightly attached to the cloud customers' data. It is extracted when the user data on the cloud's remote storage server are accessed, and is copied along with data whenever it is copied. The component deals an exceptional object or replica of the users 'current processing or accessed data and the component is accountable for the log of the exclusive access to that data object or replica. The second component logs harmonizer [8] organizes the essential centralized element which permits the end-user to access to the log records. The first component is used to automatically couple strongly with the user data wherever

in the cloud. The second component is taking the responsibility of auditing mechanism.

7. Privacy protected public data auditing in the cloud

Cloud storage service enables the user to store their huge precious processing information's, with no concern of the location, memory capacity and reliability , on the remote storage and enable the end-users to utilize the on-demand services and on-demand applications in a simple way of deliverables. Furthermore, without bothering about the need to verify its data integrity, cloud users can be capable to simply utilize the benefits of cloud facilities such as storage as like as their local storage. Enabling public auditing of the information's processed in the cloud facility is significantly serious, therefore that users can pick to an intermediary supervisors like Third party auditing to control the integrity of the data outsourced using the cloud technology and keeping themselves away from worry of data integrity. To efficiently incorporate a highly secured Third party auditing, the audit procedure should bring in no new security violations or security threads toward the privacy of user data, and establish no any supplementary online additional load or burden to the user. A secure system using cloud data depot is designed that with the support of public auditing for the user's data to preserve privacy [3] to increase the users' trust towards the user data stored in it.

To accomplish the privacy maintained public auditing [3], in the cloud storage the data exclusively incorporate the unique linear authenticator with stochastic masking or covering technique. The cloud server generates a randomness which will mask with the linear integration of sampled blocks. By the process of random mask generated on the cloud server, it is difficult or even not possible for the third party auditing to access the necessary information that to construct an accurate collection of linear equations. Therefore it is not possible to derive the user's information on the cloud, it is not a subject the total number of linear composition of the similar set of file logs can collect. But in the data correctness aspect, still the right validation of the authenticity of the block of data be evaluated in an innovative manner that apply the Homomorphic Linear Authentication proposed in [4], which is founded on the short signature strategy Boneh-Lynn-Shacham (BLS) Short Signatures Scheme [4], even with the presence of the randomness..

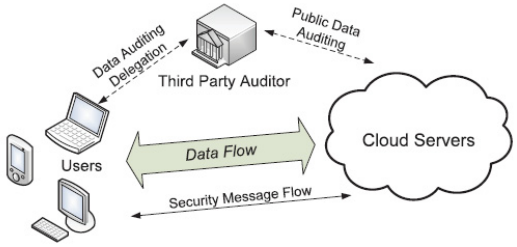

Fig. 5. The architecture of cloud data storage service [3]

8. Conclusions

Cloud computing is an extremely bright and highly potential technology that facilitates the enterprises to reduce the operational expenditures as in the point of increasing the overall performance. Although cloud computing emerged and has been distributed and utilized in the computing productivity, the security level in cloud computing is even at the level of immaturity and demands larger research focus. In this paper we have analyzed the mechanism for data accountability and auditing of cloud user data in the distribute cloud. In this context we have analyzed an innovative mechanism that technically and systematically logging any data access stored in the cloud unitedly with well supported auditing mechanism. The forward move admits the data holder in addition to inspect his information stored in the cloud but also implement well-built back-end protection in case of demand. It has also analyzed the privacy-maintaining public auditing model for user data that utilizes the cloud storage services in cloud computing with tight security. It utilize the homomorphic linear authenticator (HLA) and stochastic masking to promise their the third party audit would not able to discover any information about the user precious data or informational content stored on the data storage cloud services throughout the well-organized auditing process, that not alone reduce the load of cloud user from the dreary and highly valuable auditing mechanism, merely eliminate the cloud users' concern of their cloud stored data leakage and increase the trust the cloud.

References

[1] Chunming Rong, Son T. “Nguyen, Cloud Trends and Security Challenges”, keynote, the 3rd International Workshop on Security and Communication Networks (IWSCN 2011), May, Gjøvik, Norway.

[2] Pearson, S., "Toward Accountability in the Cloud," Internet Computing, IEEE, vol.15, no.4, pp.64, 69, July-Aug. 2011

[3] Cong Wang; Qian Wang; Kui Ren; Wenjing Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," INFOCOM, 2010 Proceedings IEEE , vol., no., pp.1,9, 14-19 March 2010

[4] Boneh, Dan, Ben Lynn, and Hovav Shacham. "Short Signatures from the Weil Pairing." Journal of Cryptology 17.4 (2004): 297-319.

[5] Radha Jagadeesan, Alan Jeffrey, Corin Pitcher, and James Riely. 2009. “Towards a theory of accountability and audit”. In Proceedings of the 14th European conference on Research in computer security (ESORICS'09), Michael Backes and Peng Ning (Eds.). Springer-Verlag, Berlin, Heidelberg, 152-167.

[6] Kailar, R., "Accountability in electronic commerce protocols," Software Engineering, IEEE Transactions on, vol.22, no.5, pp.313, 328, May 1996

[7] 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,vol. 173, pp. 187-201, 2005.

[8] Sundareswaran, S.; Squicciarini, A.C.; Lin, D., "Ensuring Distributed Accountability for Data Sharing in the Cloud," Dependable and Secure Computing, IEEE Transactions on , vol.9, no.4, pp.556,568, July-Aug. 2012

[9] "Cloud Computing Patterns." Cloud Computing Patterns. N.p., n.d. Web. 09 Feb. 2014.