Enhanced Data Sharing by Decentralized and Lightweight Framework for Accountability in Cloud Storage

Jijin Soman
Department of Pervasive Computing Technology, University College of Engineering,
BIT Campus, Tiruchirappalli.
jijinsoman@gmail.com

Abstract— Cloud computing environment provides a large scale distributed storage system where, resources on the internet are considered as a unified entity. Users use services without knowing the machines which actually process and host their own data and how computation and data storage is managed. While enjoying the convenience brought by cloud computing, the problem of losing control of user's own data from them is becoming a significant barrier to wider adoption of this emerging technology. Therefore users need a mechanism for transparent data usage of the user's data in the cloud. For achieving this goal an information accountability framework called Cloud Information Accountability is proposed in this paper which provide highly decentralized and light-weight framework to keep track of the actual usage of user's data in the cloud. CIA provides automated logging, authentication and distributed auditing mechanism to strengthen user's control over their data.

Keywords— Cloud Computing, Distributed Data Sharing, Information Accountability, Access Control, Auditing.

I. INTRODUCTION

A pool of abstracted, highly scalable, and managed compute infrastructure capable of hosting end-customer applications and billed by consumption. Amazon, Google, Microsoft, Yahoo and SalesForce are the notable CSP. The user's data are handled by huge data centers owned by CSPs. The cloud provides highly dynamic service provision and resource sharing through internet. Users are unaware of the machines which actually process their data in the cloud. This technology provides lots of advantages to users which make them conveniently use the cloud. But the level of control over the data is the current issue in this emerging technology. The control over the user's data is lost by them. The data outsourcing on clouds leads to several problems such as handling of personally identifiable information. Such issues are the reason for not adopting cloud services widely.

For overcoming these issues and satisfying users' needs an efficient technique is needed to monitor the usage of users, data in the cloud. An example scenario is ensuring whether the data are handled according to the Service Level Agreements. The approaches developed for closed domains such as databases and or centralized system won't suit for the cloud environment because of two reasons. The first reason is data handling in the cloud can be outsourced by the cloud service provider to other entities. Delegation of the tasks to others can be done by these. Second one is flexible joining and leaving of the

entities in the cloud. This makes the data handling in the cloud a complex one.

An extension of the object-oriented programming paradigm SDO is used for protecting functions or data. Sensitive functions are offered or sensitive data are hold by software objects in SDO. For maintaining safe, high-performance, and mobile code, the Proof-Carrying authentication framework can be extended. But the PCA's goal is to validate codes, rather than to monitor content.

To overcome the above problem CIA framework a novel approach, which is based on the information accountability framework is designed. A highly distributed end-to-end accountability is provided by the CIA framework which can overcome the issues in distributed data sharing in the cloud. Maintenance of lightweight and powerful accountability that combines aspects of access control, usage control and authentication is one of the featured ability of the CIA. The service level agreements can be tracked by means of the CIA. It can enforce access and usage control. The additional feature of CIA is auditing. Auditing is of two distinct modes: push mode and pull mode. Logs being periodically sent to the data owner in push mode. Users can retrieve the logs as needed in the pull mode.

The design challenges of the CIA framework are included identifying CSPs uniquely, the reliability of the log should be ensured, should adapt to a highly decentralized infrastructure, etc. These issues can be overcome by extending the programmable capability of JAR files to automatically log the usage of the users' data by any entity in the cloud. Data along with access control and logging policies are sent to cloud service providers. This is done by enclosing data and policies in JAR files. An automated and authenticated logging mechanism is triggered when accessing the data. This method is called a strong binding. Here the policies and the logging mechanism travel with the data. This mechanism allows users to the user to control his data at any location. Even though the CIA's JAR programmable capabilities meet the nature of the cloud some problems such as integrity of log files are also arises. This problem can be overcome by recording the error correction information. The Drawbacks in Existing System include,

- The complex and dynamic hierarchical service chain that handles data in the cloud is different from the conventional environments.
- It cloud is not always clear to individuals so concerns will arise from the personal information of the user are passed to other parties.
- Federated systems have problems such as end-to-end trust management and accountability.
- Only limited features are provided by the privacy manager that does not guarantee protection once the data are being disclosed.

II. Related Works

Cloud computing is Internet-based computing platform that resides in a large data centre, whereby shared resources, software and information are provided to computers and other devices on demand, like every day utility. The large data centre is managed by a third party. The third party provides computing resources as if it were a utility. The cloud environment in Fig.1 gives the cloud network where the provided services and resources can be accessed by anyone, from anywhere through the Internet. Cloud computing addresses a wide range of needs such as scientific research, e-commerce etc. One of the major problems based on **[5]** on providing computing resources as if it were a utility is Information policy. The included issues are privacy, security, reliability, access, and regulation. s service used by a great many individuals and organizations internationally, cloud computing is expanding in a rapid fashion. So policy issues related to cloud computing are not considered widely. A wide range of policy issues related to cloud computing are there that need attention. In [5] the author's introduces the policy concerns and potential solutions related to cloud computing.

The issue of how to provide appropriate privacy protection for cloud computing is important, and as yet unresolved. Pearson and Charlesworth [13] propose an approach in which procedural and technical solutions are co-designed to demonstrate accountability to resolve privacy and security risks within the cloud. In some situations Outsourcing, Offshoring, Virtualization and Autonomic technology become disadvantages in maintaining a level of privacy assurance sufficient to sustain confidence in potential customers. So authors in [13] propose Accountability concept which provides a solution to account privacy and security issues in the cloud.

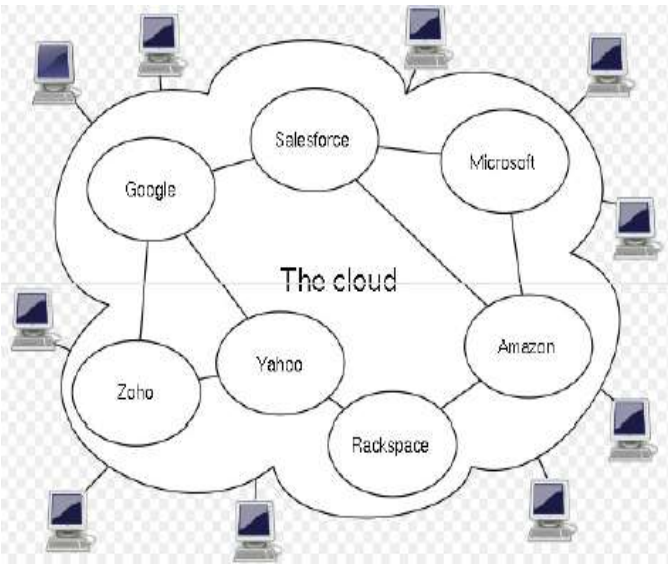

Fig.1.Cloud Environment.

Ease of information flow is an important part of large-scale, decentralized systems. Sensitive personal data leaked, corporate sensitive information revealed, copyrighted material distributed without authorization, confidential files shared among organizations in violation of regulation and policy are some of the excesses and abuses in the use of information which are viewed through information security. The existing access restrictions alone are inadequate for addressing these issues. In spite of the traditional hide it or lose it method the authors' in [17] tells about a novel mechanism called information accountability. System design is done in such a way that is oriented toward information accountability and appropriate use rather than information security and access restriction for information since the information can be more easily copied. Accountability can become a primary means which can address issues of appropriate use. The accountability concept explained in [17] is adopted for distributed and secure data sharing in this paper.

In [6] Jagadeesan, Jeffrey, Pitcher and Riely explain the foundations for distributed accountability systems. They suggested an operational model and developed analysis methods and also permits the audit based accountability systems in [6].

Remote data integrity checking is crucial in cloud computing. A Third Party Auditor (TPA) is used in [18]. With the help of a third party auditor data dynamics and public verifiability can be supported in the existing system. Through a formal analysis, the correctness and security of the protocol can be found. After that, through theoretical analysis and experimental results, the authors demonstrate that the proposed protocol has a good performance. Remote data integrity checking is done by challenging the server by the client about the integrity certain data file. The server generates responses proving that the data is complete and uncorrupted. The client should be able to verify integrity for an unlimited number of times rather than accessing the complete original data file when performing the verification of data integrity.

World wide adoption of cloud is lagging because of the customers' lack of trust in the cloud. So the authors of [1] propose a solution in their paper by providing a decentralized trust management and data accountability. The security and be comparatively replaced by cloud accountability and adaptability. The need for cloud accountability is to overcome complexity due to Virtualization and data distribution Cloud computing expect to transfer the control of computing resources partially or fully to CSPs.

The Usage Control concept in [12] constitutes of conventional access control, digital rights management, and trust management. These three are combined to provide a promising method for the next step of access control. Solution for authorizing strangers in the open environment can be done by Trust management. Controlling digital information in client-side is dealt with Digital rights management. The Usage Control (UCON) model in [12] systematically unifies these concepts.

In a paper [2] the authors explained about logic for auditing accountability in decentralized systems. Audited agents could prove their actions and authorization through logic. Data accountability and agent accountability are the constituents of accountability.

III. CIA

CIA framework is a novel approach which is based on the information accountability. In a traditional environment, the privacy protection technologies are built on the notion of hide-it-or-lose-it context. But information accountability framework provides a solution for the problem of losing control over the user's data in the cloud, by focusing on keeping the data usage transparent and trackable. A highly distributed end-to-end accountability is provided by the CIA framework which can overcome the issues. Maintenance of lightweight and powerful accountability that combines aspects of access control, usage control and authentication is one of the featured ability of the CIA. The service level agreements can be tracked by means of the CIA. It can enforce access and usage control. The additional feature of CIA is auditing. Auditing is of two distinct modes: push mode and pull mode. Logs being periodically sent to the data owner in the push mode, Users can retrieve the logs as needed in the pull mode.

The design challenges of the CIA framework are included identifying CSPs uniquely, the reliability of the log should be ensured, should adapt to a highly decentralized infrastructure, etc. These issues can be overcome by extending the programmable capability of JAR files to automatically log the usage of the users' data by any entity in the cloud. Any access performed by any entity in the cloud is automatically logging and provide a distributed auditing in the CIA framework. Logger and Log Harmonizer are the two main components in the CIA.

Users' data are strongly coupled by Logger. Logger generates log files automatically on any access to the data it coupled with. The log file is then encrypted using Identity Based Encryption. The encrypted file is sent to the Log Harmonizer. Logger also verifies the usage policies and authentication. Logger needs only minimal support from JVM. It also generates the error correction information.

Auditing is done by Log Harmonizer. The master key is generated by log harmonizer. It decrypts the log file sent by the Logger. Log Harmonizer consists of two auditory modes namely Push mode and Pull mode. These components are implemented as a portable Java Archive file. Logger is a nested Java Archive file which contains one outer Jar and one inner JAR. Outer JAR is assigned for authentication policy implementation. Inner JAR holds data and log files. Features include (1). The automatic and enforceable logging mechanism makes the data handling in the cloud easier. (2)Architecture is platform independent and highly decentralized. (3)A systematic approach to data accountability through the novel usage of JAR files. (4) The efficiency, scalability, and granularity.

IV. System Architecture

Cloud Information Accountability system architecture in fig. 2 consists of Data, Data Owner, User, Cloud Server and Logger. Each data owner registers in the cloud server and gets public and private keys based on identity based encryption. Using the generated private key data owner creates a logger component. Logger component is a nested JAR file which contains the encrypted data along with access control and authorization policies. Each data owner uploads the JAR file into the cloud server. Cloud Service Provider (CSP) allows the authorized users to access the data provided

by the data owner. For each access to the data a log file is automatically generated. The log record is encrypted using public key of the data owner thereby preventing unauthorized

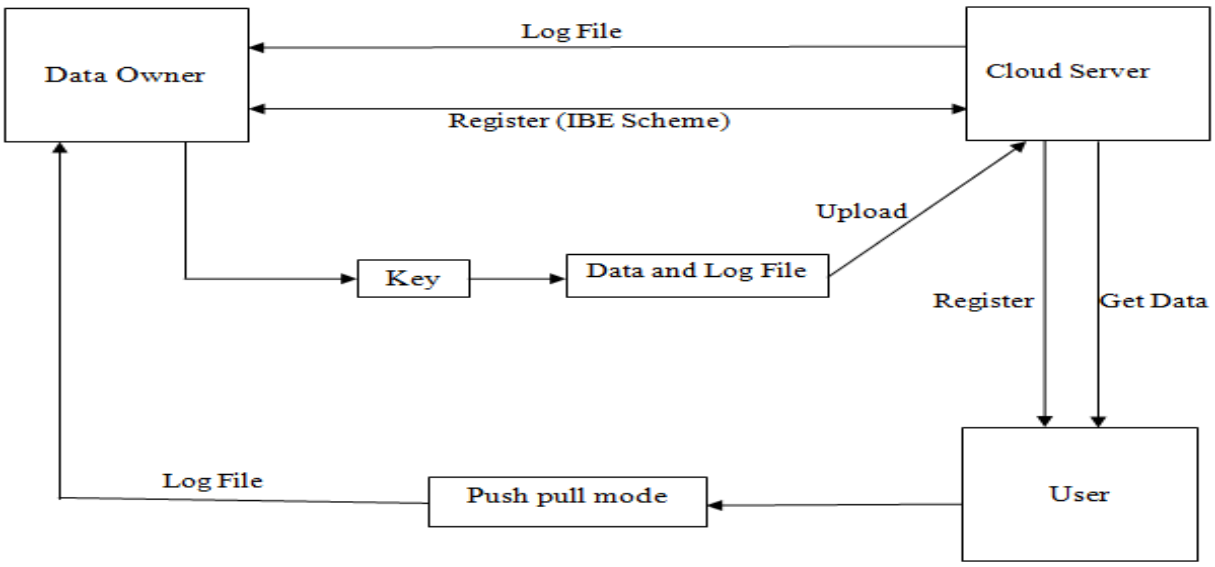

Fig.2. System Architecture

access. The encrypted log file later sends back to the data owner from both the cloud server and data used for auditing. The downloaded data file is decrypted using data owners public key by the user since obtained data would be in encrypted format. This is implemented using CSP and the data store in this paper.

This process can be explained in detail as below. Every data owner must register their details in the cloud server. And Cloud server establishes the public key and private key access policy using IBE scheme. Finally the cloud server distributes the secret key to the data owner. The cloud server stores the data owner details in the data store as an entity. The data store is a persistent storage.

After getting key's data owner creates the logger (it contains configuration details). Data owner encrypts the data using the secret key established by the cloud server. Then encrypted data into the JAR file along with access control and authorization policies. The owner data and log file are bounded or coupled together in the logger.

Every user who is purchasing the data provided by data owner must register their details and account details in the cloud server. And Cloud server establishes the public key and private key access policy using IBE scheme. Cloud server stores the user details in the data store as an entity. When the user request to get the data from the cloud server, the user gets the data with log file this log file store the user session details and finally this log file sends back to the data owner. The data owner maintains the auditing functionality. Data owner gets the log file information from the cloud server and user separately and decrypts the log files using the secret keys of data owner and then audit the log files.

V. CONCLUSIONS

CIA provides a highly decentralized light-weight accountability framework by proposing approaches for automatically logging any access to the data in the cloud. It provides an innovative auditing mechanism to the data owners along with distributed information accountability. The system is designed in such a way that it can provide a

strong back end protection if needed. Data owner can not only audit his content but also audit copies of his data that were made without his knowledge.

REFERENCES

[1] S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] B. Chun and A.C. Bavier, "Decentralized Trust Management and Accountability in Federated Systems," Proc. Ann. Hawaii Int'lConf. System Sciences (HICSS), 2004.

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

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

[5] J.W. Holford, W.J. Caelli, and A.W. Rhodes, "Using Self-Defending Objects to Develop Security Aware Applications in Java," Proc. 27th Australasian Conf. Computer Science, vol. 26,pp. 341-349, 2004.

[6] P.T. Jaeger, J. Lin, and J.M. Grimes, "Cloud Computing and Information Policy: Computing in a Policy Cloud?," J. Information Technology and Politics, vol. 5, no. 3, pp. 269-283, 2009.

[7] R. Jagadeesan, A. Jeffrey, C. Pitcher, and J. Riely, "Towards a Theory of Accountability and Audit," Proc. 14th European Conf. esearch in Computer Security (ESORICS), pp. 152-167, 2009.

[8] R. Kailar, "Accountability in Electronic Commerce Protocol s," IEEE Trans. Software Eng., vol. 22, no. 5, pp. 313-328, May 1996.

[9] W. Lee, A. Cinzia Squicciarini, and E. Bertino, "The Design and Evaluation of Accountable Grid Computing System," Proc. 29th IEEE Int'l Conf. Distributed Computing Systems (ICDCS '09), pp. 145-154, 2009.

[10] J.H. Lin, R.L. Geiger, R.R. Smith, A.W. Chan, and S. Wanchoo , Method for Authenticating a Java Archive (jar) for Portable Devices, US Patent 6,766,353, July 2004.

[11] T. Mather, S. Kumaraswamy, and S. Latif, Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance (Theory in Practice), first ed. O' Reilly, 2009.

[12] M.C. Mont, S. Pearson, and P. Bramhall, "Towards Accountable Management of Identity and Privacy: Sticky Policies and Enforceable Tracing Services," Proc. Int'l Workshop Database and Expert Systems Applications (DEXA), pp. 377-382, 2003.

[13] J. Park and R. Sandhu, "Towards Usage Control Models: Beyond Traditional Access Control," SACMAT '02: Proc. Seventh ACM Symp. Access Control Models and Technologies, pp. 57-64, 2002.

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

[15] A. Pretschner, M. Hilty, and D. Basin, "Distributed Usage Control," Comm. ACM, vol. 49, no. 9, pp. 39-44, Sept. 2006.

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

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

[18] 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.

[19] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. European Conf. Research in Computer Security (ESORICS), pp. 355-370, 2009.

[20] M. Xu, X. Jiang, R. Sandhu, and X. Zhang, "Towards a VMM Based Usage Control Framework for OS Kernel Integrity Protection," SACMAT '07: Proc. 12th ACM Symp. Access Control Models and Technologies, pp. 71-80, 2007.