Accountability for Data Sharing in the Cloud

Anuja Musale[1], Chaitra Rane[2], Nikita Zambad[3], Anuja Birari[4]

[1, 2. 3, 4]Student, Bharatividyapeeth College of Engineering for Women, Pune University, Katraj, Pune-043, India

Abstract: ***Cloud computing enables highly scalable services to be easily consumed over the Internet on an as-needed basis. A major feature of the cloud services is that users' data are usually processed remotely in unknown machines that users do not own or operate. While enjoying the convenience brought by this new emerging technology, users' fears of losing control of their own data (particularly, financial and health data) can become a significant barrier to the wide adoption of cloud services. To address this problem, in this paper, we propose a novel highly decentralized information accountability framework to keep track of the actual usage of the users' data in the cloud. In particular, we propose an object-centered approach that enables enclosing our logging mechanism together with users' data and policies. We leverage the JAR programmable capabilities to both create a dynamic and traveling object, and to ensure that any access to users' data will trigger authentication and automated logging local to the JARs.***

Keywords: cloud computing, accountability, data sharing, information accountability framework

1. Introduction

Cloud computing presents a way to supplement the current consumption and delivery model for IT services based on the Internet, by providing for dynamically scalable and often virtualized resources as a service over the Internet. To date, there are a number of notable commercial and individual cloud computing services, including Amazon, Google, Microsoft, Yahoo, and sales force. While users enjoy the convenience brought about by this new emerging technology, users' also fear of losing control of their own data (particularly, financial and health data) and this can become a significant barrier to the wide adoption of cloud services.

To reduce 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. Conventional access control approaches developed for closed domains such as databases and operating systems, or approaches using a centralized server in distributed environments, are not suitable, due to the following features characterizing cloud environments. First, data handling can be outsourced by the direct cloud service provider (CSP) to other entities in the cloud and theses entities can also delegate the tasks to others, and so on. Second, entities are allowed to join and leave the cloud in a flexible manner. As a result, data handling in the cloud goes through a complex and dynamic hierarchical service chain which does not exist in conventional environments.

To overcome the above problems, we propose a novel approach, namely Cloud Information Accountability (CIA) framework, based on the notion of information accountability. Unlike privacy protection technologies which are built on the hide-it-or-lose-it perspective, information accountability focuses on keeping the data usage transparent and trackable. Our proposed CIA framework provides end-to-end accountability in a highly distributed fashion. One of the main innovative features of the CIA framework lies in its ability of maintaining lightweight and powerful accountability that combines aspects of access control, usage control and authentication. By means of the CIA, data owners can track not only whether or not the service-level agreements are being honored, but also enforce access and usage control rules as needed. Associated with the accountability feature, we also develop two distinct modes for auditing: push mode and pull mode. The push mode refers to logs being periodically sent to the data owner or stakeholder while the pull mode refers to an alternative approach whereby the user (or another authorized party) can retrieve the logs as needed.

2. Problem Statement

To propose a novel highly decentralized information accountability framework to keep track of the actual usage of the users' data in the cloud that enables enclosing our logging mechanism together with users' data and policies to ensure that any access to users' data will trigger authentication and automated logging local to the JARs.

3. Related Work

Smitha Sundareswaran, Anna C. Squicciarini, and Dan Lin, [1] provide innovative approaches for automatically logging any access to the data in the cloud together with an auditing mechanism. Their approach also allows the data owner to not only audit his content but also enforce strong back-end protection if needed. But they fail to provide hashing of log files for faster retrieval of logs. R. Corin, S. Etalle, J.I. den Hartog, G. Lenzini, and I. Staicu [8] proposes a language that allows agents to distribute data with usage policies in a decentralized architecture. This paper designs a logic that allows audited agents to prove their actions, and to prove their authorization to possess particular data. This paper gives a general idea about attaching usage and access policies to the data of the data owner.

4. Proposed System

In this paper, we propose a novel highly decentralized information accountability framework to keep track of the actual usage of the users' data in the cloud. In particular, we propose an object-centered approach that enables enclosing our logging mechanism together with users' data and policies. We leverage the JAR programmable capabilities to both create a dynamic and traveling object, and to ensure that

any access to users' data will trigger authentication and automated logging local to the JARs. To strengthen user's control, we also provide distributed auditing mechanisms. We provide extensive experimental studies that demonstrate the efficiency and effectiveness of the proposed approaches.

5. Approach

We propose innovative approaches for automatically logging any access to the data in the cloud together with an auditing mechanism. Our approach allows the data owner to not only audit his content but also enforce strong back-end protection if needed. Moreover, one of the main features of our work is that it enables the data owner to audit even those copies of its data that were made without his knowledge. In the future, we plan to refine our approach to verify the integrity of the JRE and the authentication of JARs. For example, we will investigate whether it is possible to leverage the notion of a secure JVM [18] being developed by IBM. This research is aimed at providing software tamper resistance to work required in the aftermath of a database corruption. This saves both time and money for those affected. The techniques also highlight the advantages over approaches relying heavily on information restriction through either hardware which can have prohibitive costs for small institutions, have a limited shelf-life and are relatively complex; or cryptography which does not adequately offer remedies after a leak.

5.1 Usefulness/ Advantages

We propose innovative approaches for automatically logging any access to the data in the cloud together with an auditing mechanism. Our approach allows the data owner to not only audit his content but also enforce strong back-end protection if needed. Moreover, one of the main Features of our work are that it enables the data owner to audit even those copies of its data that were made without this knowledge.

6. Figures and Tables

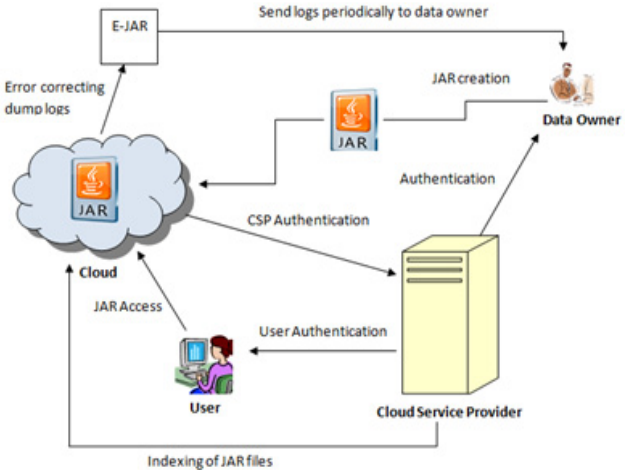

Figure 5.1

The above image provides a basic idea of the system being developed.

7. Assumption and Dependencies

1. We are going to use Private cloud in this project.
   a. Private cloud: Here we are using 3-4 machines to form a cloud. One machine will be the server and others will be client. Client will send the request and server will process that request
2. Before doing the setup on cloud we have to test our project on server then after checking its functionality we will create setup on cloud.
3. Here we are assuming that one machine will be server and others will be client.

8. Future Scope

In the future, we plan to refine our approach to verify the integrity of the JRE and the authentication of JARs. For example, we will investigate whether it is possible to leverage the notion of a security. This research is aimed at providing software tamper resistance to Java applications. In the long term, we plan to design a comprehensive and more generic object-oriented approach to facilitate autonomous protection of traveling content. We would like to support a variety of security policies, like indexing policies for text files, usage control for executables, and generic accountability and provenance controls.

9. Conclusion

We introduced modern approaches for automatically logging any access to the data in the cloud together with an auditing mechanism. Our approach allows the data owner to not only audit his content but also enforce strong back-end protection if needed. Apart from that we have enclosed methodology to enhance the integrity of owner's data. In future, we plan to refine our approach to verify the integrity of JRE. For that we will look into whether it is possible to leverage the advantage of secure JVM being developed by IBM and we would like to enhance our architecture from user end which will allow the users to check data remotely in an efficient manner in multi cloud environment

References

[1] SmithaSundareswaran, Anna C. Squicciarini, and Dan Lin, "Ensuring Distributed Accountability for Data Sharing in the Cloud", IEEE Transactions On Dependable and Secure Computing, Vol. 9, No. 4, July/August 2012.
[2] Distributed Usage Control_Alexander Pretschner, Manuel Hilty, David Basin, Computer systems play an increasingly prominent role in our daily lives. Interacting with these systems often involves disclosing personal data, i.e., data that can be traced back to particular individuals, collected in different contexts. The looming reality of ubiquitous computing will further increase the amount of personal data collected, and enhanced network capabilities give rise to potentially uncontrolled distribution. These technologies improve, for the most part, the quality of our lives. Still, the question arises how all this potentially sensitive data can be protected. Two of the main technical challenges here are controlling data access and usage. While the fundamentals of access control appear to be well understood, this is not the case for usage control. Promising research has been carried out in the areas of both usage control specification and enforcement mechanisms. Missing though is a conceptual

framework that encompasses both specification and enforcement. To this end, we assume that personal and other kinds of sensitive data are stored at trustworthy places called data providers. Third parties, called data consumers, request access to the data. Assuming that some form of access control is in place, our concern is what happens to the data once it has been released to the data consumer, i.e., how the data consumer may, must, and must not use it. Clearly, the scope of this problem extends beyond privacy concerns about personal data and is also related to the management of intellectual property rights.

[3] A Logic for Auditing Accountability in Decentralized System R. Corin1, S. Etalle1;2, J. den Hartog1, G. Lenzini1, and I. Staicu1

[4] This paper proposes a language that allows agents to distribute data with usage policies in a decentralized architecture. In the given framework, the compliance with usage policies is not enforced. However, agents may be audited by an authority at an arbitrary moment in time. This paper designs a logic that allows audited agents to prove their actions, and to prove their authorization to posses particular data. Accountability is defined in several flavors, including agent accountability and data accountability. This paper presents a logic data access and agent accountability in a setting in which data can be created, distributed and re-distributed. Using this logic, the owner of the data attaches a usage policy to the data, which contains a logical specification of what actions are allowed with the data, and under which conditions. This logic allows for different kind of accountability and it is shown to be sound.

[5] Decentralized Trust Management and Accountability in Federated Systems Brent N. Chun (Intel Research Berkeley), Andy Bavier(Princeton University) This paper describes three key problems for trust management in federated systems and present a layered architecture for addressing them. The three problems addressed include how to express and verify trust in a flexible and scalable manner, how to monitor the use of trust relationships over time, and how to manage and re-evaluate trust relationships based on historical traces of past behavior. While previous work provides the basis for expressing and verifying trust, it does not address the concurrent problems of how to continuously monitor and manage trust relationships over time. These problems close the loop on trust management and are especially relevant in the context of federated systems where remote resources can be acquired across multiple administrative domains and used in potentially undesirable ways (e.g., to launch denial-of-service attacks)

[6] Identity-Based Encryption from the Weil Pairing Dan Boneh_ Matthew Frankliny Appears in SIAM J. of Computing, Vol. 32, No. 3, pp. 586-615, 2003. An extended abstract of this paper appears in the Proceedings of Crypto 2001, volume 2139 of Lecture Notes in Computer Science, pages 213{229, Springer-Verlag, 2001. The paper proposes a fully functional identity-based encryption scheme (IBE). The scheme has chosen cipher text security in the random oracle model assuming a variant of the computational Diffie-Hellman problem. This system is based on bilinear maps between groups. The Weil pairing on elliptic curves is an example of such a map. This paper gives precise definitions for secure identity based encryption schemes and gives several applications for such systems.

[7] A Privacy Manager for Cloud Computing Siani Pearson, Yun Shen and Miranda Mowbray HP Labs, Long Down Avenue, Stoke Gifford, Bristol BS34 8QZ, UK This paper describes a privacy manager for cloud computing, which reduces the risk to the cloud computing user of their private data being stolen or misused, and also assists the cloud computing provider to conform to privacy law. It also describes different possible architectures for privacy management in cloud computing; give an algebraic description of obfuscation, one of the features of the privacy manager; and describe how the privacy manager might be used to protect private metadata of online photos.

Author Profile

Anuja Musale is a final year computer engineering student of Bharati Vidyapeeth College of engineering for women. The college is affiliated to Pune University.

Chaitra Raneis a final year computer engineering student of Bharati Vidyapeeth College of engineering for women. The college is affiliated to Pune University.

Nikita Zambad is a final year computer engineering student of Bharati Vidyapeeth College of engineering for women. The college is affiliated to Pune University.

Anuja Birari is a final year computer engineering student of Bharati Vidyapeeth College of engineering for women. The college is affiliated to Pune University.