Make sure Appropriate Responsibility for Data Sharing and Data Integrity in the Cloud

Rahul D. Kuwar
M. Tech Student, Dept. of Computer Engineering, AVN Institute of Engineering & Technology, Hyderabad, TS, India

B. Pannalal
Asst. Professor, Dept. of Computer Engineering, AVN Institute of Engineering & Technology, Hyderabad, TS, India

Dr. Shaik Abdul Nabi
HOD, Dept. of Computer Engineering, AVN Institute of Engineering & Technology, Hyderabad, TS, India

Abstract – Distributed computing is the utilization of assets like equipment and programming that are conveyed as an administration over the system/web. In today numerous associations, government, people utilized distributed computing. With the utilization distributed computing you to get to your data from anyplace whenever. While a typical PC setup expects you to be in an indistinguishable area from your information stockpiling gadget, the cloud takes away that procedure. The distributed computing expels the requirement for you to be in an indistinguishable area from the equipment which is stores your information. A noteworthy element of the cloud administrations is that client's information is normally prepared remotely in obscure machines that clients don't possess or work. Information dealing with in the cloud experiences an unpredictable and dynamic various levelled benefit chain which does not exist in current situations. Information taking care of in the cloud experiences a mind boggling and dynamic various levelled benefit chain which does not exist in existing situations. This can recognize by in particular Cloud Information Accountability (CIA) system. In this each and each entrance to the information are effectively and consequently logged. Log records ought to be sent back to their information proprietors or administrator intermittently to educate them of the present use of their information. All the more significantly, client log documents ought to be retrievable whenever by their information proprietors when required in any case the area where the records are put away. The end client is permitted to get to the information according to their client/get to rights which they indicate while enrolling to get to the information in the cloud and verification is given by administrator and information client will be checked. To reinforce client's control access of information in the distributed computing administration, a novel approach, to be specific cloud data responsibility (CIA) system is proposed which secure information amid transmission and capacity. Trial comes about are accounted for to exhibit the proficiency and adequacy of the proposed approach.

Index Terms – accountability, cloud computing, data sharing, data integrity

1. INTRODUCTION

In a world that sees new technological trends bloom and dull on almost a daily basis, one new trend promises more longevity. This trend is called cloud computing, and it will change the way you use your computer and the Internet. Cloud computing project a major change in how we store information and run applications. Instead of running programs and data on an individual computer, everything is hosted in the "cloud"—a nebulous assemblage of computers and servers accessed via the Internet. In Cloud computing you access all your applications and documents from anywhere in the world. Ordinary access control approaches produced for shut spaces, for example, databases and working frameworks, or methodologies utilizing a concentrated server in circulated situations, are not appropriate, because of the accompanying highlights describing cloud conditions. To begin with, information taking care of can be outsourced by the immediate cloud specialist organization (CSP) to different substances in the cloud and theories elements can likewise appoint the errands to others, et cetera. Second, substances are permitted to join and leave the cloud in an adaptable way. Therefore, information taking care of in the cloud experiences an intricate and dynamic various levelled benefit chain which does not exist in regular situations. To conquer the above issues, we propose a novel approach, specifically Cloud Information Accountability (CIA) system, in view of the idea of data responsibility. Dissimilar to security insurance advances which are based on the shroud it-or-lose-it point of view, data account-capacity concentrates on keeping the information use straightforward and track capable. Distributed

Computing is the utilization of processing over a system (ordinarily the web). Distributed computing alludes to the conveyance of processing and capacity limit as an administration to a heterogeneous group of end clients the name originates from the utilization of mists as a reflection for the complex structure. It gives remote administrations a client's information, programming and calculation over a system. Distributed computing is the new term for the since quite a while ago imagined vision of registering as a utility. Distributed computing is a stage that lives in a major server farm and can powerfully give servers the capacity to address an extensive variety of necessities, going from logical research to internet business. Distributed computing is extending quickly as administration utilized by a many single and associations globally, arrangement issues identified with distributed computing. Points of interest of the administrations gave are disconnected from the clients who never again should be specialists of innovation foundation.

2. EXISTING SYSTEM

To mollify clients' worries, it is basic to give a powerful instrument to clients to screen the use of their information in the cloud. For instance, clients should have the capacity to guarantee that their information are taken care of as per the administration level assertions set aside a few minutes they sign on for administrations in the cloud. Regular access control approaches created for shut spaces, for example, databases and working frameworks, or methodologies utilizing a brought together server in circulated conditions, are not appropriate, because of the accompanying highlights portraying cloud situations.

2.1 PROBLEMS ON EXISTING SYSTEM

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. Third, although the Cloud computing is vast developing technology, the database management system does not have a trustworthiness.

3. PORPOSED SYSTEM

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 track able. 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.

Proposed System to overcome the above problems, we propose a novel method, namely Cloud Information Accountability (CIA) framework, based on the notion of information accountability. Data Owner can upload the data into the cloud server after encrypted the data. User can subscribe into the cloud server with certain access polices such as read, write and copy of the original data. The Loggers and Log Harmonizer will have a track of the access logs and reports to the data owner. This Process ensures security.

Our main contributions are as follows:

•We propose a novel automatic and enforceable logging mechanism in the cloud.

•Our proposed architecture is platform independent and highly decentralized, in that it does not require any dedicated authentication or storage system in place.

•We go beyond traditional access control in that we provide a certain degree of usage control for the protected data after these are delivered to the receiver.

•We conduct experiments on a real cloud testbed. The results demonstrate the efficiency, scalability, and granularity of our approach. We also provide a detailed security analysis and discuss the reliability and strength of our architecture.

•We can share the data in a secured manner.

4. ARCHITECTURE

The figure 4.1 Shows the architecture of the proposed system and this architecture is same explained and also applicable to our proposed approach. The overall CIA framework, combining data, users, logger and harmonizer is sketched in Fig. 4.1. At the beginning, each user creates a pair of public and private keys based on Identity-Based Encryption (step 1 in Fig. 4.1). This IBE scheme is a wail- pairing- based IBE scheme, 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 stakeholders (users, companies) are authorized to access the content itself. Then, he sends the JAR file to the cloud service provider that he subscribes to. To authenticate the CSP to the JAR (steps 3-5), use OpenSSL-based certificates, wherein a trusted certificate authority certifies the CSP. In the event that the access is requested by a user employ SAML-based authentication, where in a trusted identity provider issues certificates verifying the user's identity based on his username. Once the authentication succeeds, the service provider (or the user) will be allowed to access the data enclosed in the JAR. Depending on the configuration settings defined at the time of creation, the JAR will provide usage control associated with logging, or will provide only logging functionality.

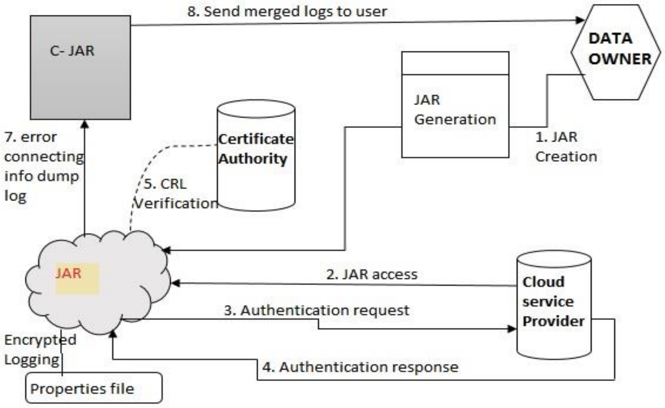

Fig 4.1: System Architecture

As for the logging, each time there is an access to the data, the JAR will automatically generate a log record, encrypt it using the public key distributed by the data owner, and store it along with the data (step 6). The encryption of the log file prevents unauthorized changes to the file by attackers. The data owner could opt to reuse the same key pair for all JARs or create different key pairs for separate JARs. Using separate keys can enhance the security without introducing any overhead except in the initialization phase. In addition, some error correction information will be sent to the log harmonizer to handle. The concept of the cloud computing model is that customers' data, which can be of individuals, organizations or enterprises, is processed remotely in unknown machines that users do not own or operate. The convenience and efficiency of this approach, however, comes with privacy and security risks. A considerable obstacle to the adoption of cloud services is the users' fear of confidential data leakage and loss of privacy in the cloud. The process of protection of users' data begins from the stage the user starts his cloud experience. The user has to manually identify the cloud provider that meets his privacy requirements, and this is often significant burden for end-users. Also many more problems arise after storing data on the clouds. With respect to the above mentioned scenario we can identify some common requirements needed to achieve data accountability in the cloud. With the above scenario in mind, we identify the common requirements and develop several guidelines to achieve data accountability in the cloud. A user who subscribed to a certain cloud service, usually needs to send his/her data as well as associated access control policies (if any) to the service provider. After the data are received by the cloud service provider, the service provider will have granted access rights, such as read, write, and copy, on the data. Using conventional access control mechanisms, once the access rights are granted, the data will be fully available at the service provider. In order to track the actual usage of the data, we need to develop logging and auditing techniques.

5. MODULE DESCRIPTION

5.1 User/Data Owner

User is the person is going to see or download the data from the Cloud server. To access the data from the Cloud server, the users have to be registered with the cloud server. So that the user has to register their details like username, password and a set of random numbers. This is the information that will store in the database for the future authentication. Data Owner is the Person who is going to upload the data in the Cloud Server. In order to upload the data into the Cloud server, the Data Owner have to be registered in the Cloud Server. Once the Data Owner registered in cloud server, the space will be assigned to the Data Owner.

5.2 Cloud Server

Cloud Server is the area where the user going to request the data and also the data owner will upload their data. Once the user send the request regarding the data they want, the request will be first send to the Cloud Server and the Cloud Server will forward your request to the data owner. The data Owner will send the data the data the user via Cloud Server. The Cloud Server will also manage the Data owner and Users information in their Database for future purpose.

5.3 Logger

The Logger is maintained by the Cloud Server. Loggers have the details of the data owner and users who are accessing the Cloud Server. So the Logger will be more useful for many purposes. Like which user / data owner accessing the Cloud Server, accessed at the particular time and the IP address from which the data is requested by user etc.

5.4 Certificate Authority

The Certificate Authority is utilized to check the Cloud Server is perceived or not. The Cloud Server must be perceived by the endorsement specialist. If not perceived, the Cloud Server is a Fraudulent Server. The information proprietor can check the whether the perceived or not. Since the information proprietor will transfer their information in the Cloud Server.

5.5 Access Privileges

The access privileges are set by the data owner for accessing their data. Some Owners will provide read only, some of them will allow read and download. The Cloud Server will send the dynamic intimation when the user is accessing the data beyond their limits. This increases more security while sharing the data in the Cloud.

5.6 Push and Pull Concept

Push

For every periodical time the Cloud Server will send the access details of the user to the data owner. So that the Data Owner may able to know who're all the accessing their data at the particular time period. During the registration phase, the Data owner will ask by the Cloud Server whether they're choosing the push or pull method.

Pull

In the Pull method, the data owner has to send the request to the Cloud Server regarding the access details of their data up to the particular time. Then the Cloud Server will send the response to the Data Owner regarding the user's access details.

5.7 Random Set Generation and Verification

When the user requests the data to be downloaded from the Cloud Server, the user have to enter the Random number set. If it is matched, the user is allowed to download the data. The Random number sets will be providing to the user during the registration Phase itself. Each and Every time the Random number set will vary. This ensures security while downloading the data.

6. CONCLUSION AND FUTURE SCOPE

Conclusion part depicts the main points as the constructive finds obtained from the proposed system. Conclusion should not be the same as abstract. Conclusion should be modelled efficiently Cloud computing has raised a range of important privacy and security issues. Such issues are due to the fact that, in the cloud, users' data and applications reside at least for a certain amount of time on the cloud cluster which is owned and maintained by a third party. Concerns arise since in the cloud it is not always clear to individuals why their personal information is requested or how it will be used or passed on to other parties. The privacy problem in the cloud is also compounded by the fact that some of the issues are not technical in nature, and rather deal with law and regulations. To achieve accountability along with security and privacy of data on the cloud we have developed a system based on Cloud Information Accountability Framework. This approach uses the programmable capability of JAR (Java Archives) files to automatically log the usage of the users' data by any entity in the cloud. Users send their data along with any policies such as access control policies and logging policies that they want to enforce, enclosed in JAR files, to cloud service providers. Any access to the data triggers an automated and authenticated logging mechanism local to the JARs. The data is encrypted and stored in the inner JAR. The inner JAR along with the access policies to access the data are stored in an outer JAR. The outer JAR is locked and the entire process of

Uploading and downloading data is carried out on a secure environment using OpenSSL. And the most important feature is that it enables the data owner to audit even those copies of its data that were made without his knowledge. Meta-data of the files is stored in a database, which helps the users to search for the data they need to view. The project has been developed to work on image data files. It can be further modified to work on any type of file, making it useful for any educational institutes, universities, etc. to share timetables, notices, subject notes, etc. with the students and staff. The project can be enhanced further to verify the integrity of the JRE and the authentication of JARs. Also the project can be expanded to more than on CSPs, where the data owners choose to their store data on servers of more than one CSP.

REFERENCES

[1] Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online by Michael Millar.

[2] Cloud Computing: A Practical Approach by Anthony T. Velte, Toby J. Velte, Ph.D.,RobertElsenpeter.

[3] Smitha S, Anna C.S and Dan L, ‟Ensuring distributed accountability for data in the cloud", IEEE Transactions on Dependable and Secure Computing, Vol. 9, No. 4, pp.556-568,2012.

[4] Amandeep kaur And Satinderpal Singh, "Design And Development Of Novel Distributed Information Monitoring Framework To Check Actual Information Usage Over Cloud‟, International Journal Of Innovaive Research And Development, Vol 2 Issue 8 August 2013

[5] Bheemeshwar Yerra, Amjan Shaik And M.Sudhir Kumar, „Liability As An Approach For Confidentiality Fortification In The Cloud‟, International Journal Of Computer And Electronics Research, Volume 2, Issue 4, August 2013.

[6] Boneh D and Franklin M.K, „Identity-based encryption from the well pairing‟, Proc Int'l Cryptography Conf. Advances in Cryptography, pp. 213-229, 2012 .

[7] Buneman P, Chapman A and Cheney J, „Provenance management in crated databases., Proc. ACM SIGMOD Int'l Conf. Management of Data(SIGMOD'06), PP. 539-550,2006.

[8] Epuru Madhavrao, M Parimala And Chikala Jaya Raju, " Data Sharing in the Cloud Using Distributed Accountability‟, International Journal of Advanced Research in Computer Engineering & Technology, ISSN: 2278 – 1323 Volume 2, Issue 4, April 2013.

[9] Jaeger P.T, Lin J and Grimes J.M, „Cloud computing and information policy:computing in a policy cloud?‟.Information Technology and Policies, Vol. 5, No. 3, pp.269-283, 2009.

[10] Mr. R. Karthik Ganesh And Ms. Aranya Hari, „Enhancing Privacy In Cloud By . Avoiding Misuses Of Files‟, International Journal Of Advanced Research In Computer Science And Software Engineering, Volume 3, Issue 3, March 2013.K. Elissa, "Title of paper if known," unpublished.

Authors

Mr. Rahul Dilip Kuwar Completed Bachelor of Engineering from University of Pune, India and he is currently pursuing M.Tech in the stream of Computer Science and Engineering, AVN Institute of Engineering & Technology, Ibrahimpatnam, Hyderabad, TS, India. (Presented a paper on) His areas of interest are DATABASE, PHP, JAVA and Cloud computing.

B. PANNALAL is working as Assistant Professor in Dept. of CSE, AVN Institute of Engineering & Technology, Hyderabad, T.S, India. He completed his B.Tech (Computer Science & Engineering) from JNTU, Hyderabad. He has completed his M.Tech from JNTU Ananthapur campus, India. He is a certified professional in Teaching by National Institute Of Technical Teachers Training & Research (Govt Of India)

He is having 10 years of Teaching Experience in various Engineering Colleges. His expertise areas are Design and Analysis of Algorithms, Data Structures & Linux Networking Programming.

Dr. Shaik Abdul Nabi is working as professor & Head of the Dept. of CSE, AVN Inst.Of Engg.& Tech, Hyderabad, T.S, India. He completed his B.E (Computer Science) from Osmania University, Hyderabad. He has completed his M.Tech. from JNTU Hyderabad campus and he received Doctor of Philosophy (Ph.D) in the area of Web Mining from AcharyaNagarjuna University, Guntur, AP, India. He is a certified professional by Microsoft.

He is having 17 years of Teaching Experience in various Engineering Colleges. He has published 15 publications in International / National Journals and presented 08 papers in National / International conferences. His expertise areas are Data warehousing and Data Mining, Data Structures & UNIX Networking Programming, Cloud Computing and Mobile Computing.