Survey on Data Sharing In the Cloud Using Distributed Accountability

Miss. Priyanka More
PG Student, Dept. of Computer Engineering
G.H. Raisoni College Of Engineering & Management
Pune, India

Miss. Mayuri Lingayat
Assistant Professor Dept. of Computer Engineering
G.H. Raisoni College Of Engineering & Management
Pune,India

Abstract- In cloud computing, resources are share among multiple clients and it is important for cloud service provider to allocate these resources for such clients. Cloud computing is an infrastructure that provides on demand network services. In this technology, user fears about loss of his own personal data. To solve this problem, we propose data accountability approach which is used to keep track of usage of user's data in the cloud. We create JAR programmable files to ensure user's data authentication and automated log in JARs. Using this mechanism owner may know his data is handled as per his requirement only. By auditing method, we can ensure that the information which was stored in the cloud is used by the responsible persons as per the service level agreements. In this paper, this distributed accountability concept is applied to generic file types which contain not only image data type but also all types of data. It also supports different variety of security policies, like indexing policies for text files, usage control policies for executable files. It describes an application, which maintains log records of data usage stored on cloud. It also provides strong protection to owner's data by compiling jar of that file.

Keywords: Cloud computing, Security, JARs, Encryption, Accountability, Data sharing.

I. INTRODUCTION

Cloud computing is internet-based computing which contains large groups of remote servers that are interconnected to allow the centralized data storage as well as online access to various services or resources. Popularity of cloud computing is increasing rapidly in distributed computing environment. In this, user stores his own data on cloud server and accesses it through internet. The actual location of data server is unknown to the user. Basically there are three building blocks of cloud computing, software as a service (SaaS), platform as a service (PaaS) and Infrastructure as a service (Iaas), all these allow user to store their data online. Large access to data, application, resources and hardware without installing any software is one of the main feature of cloud computing. User can access the data from anywhere in the world .Data storage on cloud server shown in following Fig.1.By using this architecture data is share on cloud.

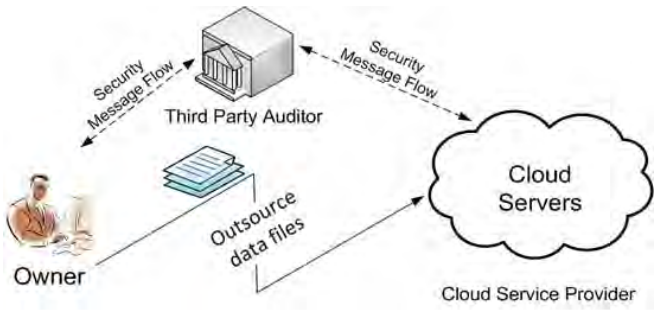

Figure1: Data storage on cloud

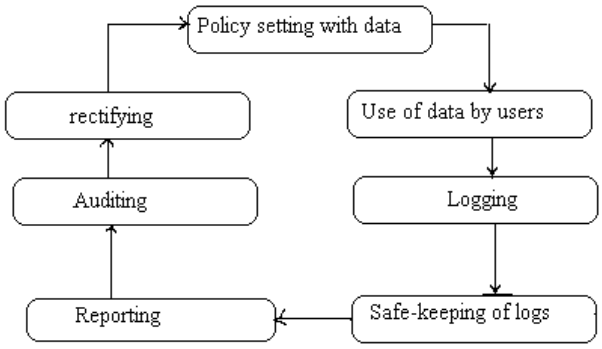

Figure.2: Phases of Accountability

Data usage in cloud is very large by users and businesses, so data security in cloud is major risk which is very important issue to solve. It is very important to solve such security related issues which are present in cloud; hence it requires strong security mechanism which will track usage of users data in the cloud. It means by using such mechanism other user cannot read the respective owners data without having access so that Data owner should not bother about his data, and should not worrying about destruction of his data by hacker.

Accountability is one of the mechanism which is necessary for monitoring usage of cloud data. It provides customers with transparency as well as control over data in the cloud. This framework also provides reliable information about usage of data and it also observes all the records. Basically accountability is for verification of authorization and authentication. It is powerful tool to check the Policies of authorization. In our distributed accountability and auditing framework, owner of data will set the policies for the data first, which he wants to place in cloud and send it to cloud service provider enclosed in JAR files, any access to such data will be automatically checked for its authentication and logs record for each data item will be created and periodically sent to data owner for monitoring the data usage. Phases of this accountability framework shown in fig. 2.

In first phase, policies are set with data by owner. According to these policies, user can use such data. In second phase this data is used by authorized user after policy setting phase only. After that whenever users use such data that time log information will save in log file. After logging phase, we need to protect the integrity of the logs to prevent unauthorized access. Here to protect the logs encryption may be applied. Next in reporting phase, Reporting tools generate from logs file-centric summaries and reports of the audit trails, access history of files. Reports and logs are checked here in auditing phase .The checking can be performed by auditors. In rectifying phase, Problem areas in the cloud are removed or rectified and governance of such cloud processes will improve.

II. LITERATURE SURVEY

In this section we review some related works which addresses security related issues in cloud data storage. Such security issues are very important in cloud. There are many techniques which are already available so this section gives the review of all these techniques.

A. Title: “Ensuring Distributed Accountability for Data Sharing in the Cloud”.

Author: Smitha Sundareswaran, Dan Lin,”IEEE Transaction on dependable a secure computing, 2012”

Summary: The authors describe automatic logging method in the cloud system. They define systematic approach to information accountability through the usage of JAR files. Their approach allows the data owner to not only audit his content but also enforce strong back-end protection if needed. Their proposed architecture is platform independent, which does not require any storage system or authentication in place. They define certificate authority in the cloud server to confirm the cloud server again. When any cloud server is not responding then such server is called as fraud server. When data owner is stored his contents on the cloud at that time he first checks the cloud server. But in this paper they were not provide hashing of log files which is necessary for faster retrieval of logs.

B. Title: “Accountability as a Way Forward for Privacy Protection in the Cloud”

Author: S. Pearson, A. Charlesworth, “Proc First Int'l conf. Cloud Computing, 2009”.

Summary: The issue of how to provide strong privacy protection for cloud computing is very important. In this paper Authors describe solution to accountability for solving risks which are related to cloud security. In this mechanism different policies are decided only by the parties that store, use or share data irrespective of the jurisdiction in which information is processed. They propose an approach in which procedural as well as technical solutions are co-designed to demonstrate accountability as a way forward to resolving security as well as privacy risks. But it has one limitation that data processed on SP is in unencrypted format at the time of processing so there is a risk of data leakage.

C. Title: “A privacy Manager for Cloud Computing”

Author: Y. Shen, M. Mowbray," “Proc. Int'l Conf. Cloud Computing, 2009”

Summary: Mowbray describes privacy manager mechanism in which user's data is safe on cloud. In this approach the users data which is in encrypted form in cloud and evaluation is done on encrypted data only, the privacy manager make readable data from result of such evaluation manager to get the correct result. In obfuscation users data is not present on Service provider's machine so there is no risk with their data, so data is defiantly safe on cloud. They describe different architectures for privacy management in cloud computing; give an algebraic description of obfuscation and also describe how the privacy manager used to protect secret metadata of online photos. But this solution is not suitable for all cloud applications, when input data is very large then this method can still require a large amount of memory.

D. Title: “Enabling public auditability and data dynamics for storages security in cloud computing”

Author: Q Wang, C. Wang, K. Ren, “INFOCOM, IEEE, 2010”

Summary: The authors proposed a dynamic auditing Protocol. This protocol supports the dynamic operations of the user’s data on the cloud servers. Their work studies the problem of ensuring the integrity of data storage in Cloud Computing. They consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the dynamic data which is stored on the cloud. TPA eliminates the involvement of the client through the auditing of whether his data which is stored on cloud is indeed intact, which can be important in achieving economies of scale for Cloud Computing. While their prior works on ensuring remote data integrity often lacks the support of either public auditability or dynamic data operations, this paper achieves both things. But this mechanism may leak the data content to the auditor because it requires the server to send the linear combinations of data blocks to the auditor.

E. Title: “A Logic for Auditing Accountability in Decentralized Systems,”

Author: R. Corin, S. Etalle, J.I. den Hartog, “Proc. Workshop Formal Aspects in Security and Trust, 2005.

Summary: The author provides a language which allows to provide data with usage policies by agent. They design a logic that allows agents to prove their authorization and action to use particular data. In this technique data owner add Policies with data that contain a detail description of which actions are allowed with which data, in their framework, the compliance with usage policies is not enforced. However, agents maybe audited by authority at arbitrary moment in time. But there is the problem of Continuous auditing of agent. But there work has also provided solution that monitors incorrect behavior of agent and agent has to give justification for their actions, after that authority will check the justification.

III. OVERVIEW OF PROPOSED IDEA

As we know that, cloud computing provides on demand services. But the data operated on clouds lead to a number of issues related to accountability, including the management of personally identifiable information. So it is necessary to provide a mechanism which effectively monitors all the usage of the data in cloud. There are multiple techniques already available as we discuss in literature survey but they have some limitations and disadvantages. So to overcome all this problems, Cloud Information Accountability (CIA) framework is introduced here which will use in our proposed idea. Purpose is that all data can be securely share on cloud.

A. Cloud Information Accountability And Auditability:

CIA framework is based on data accountability concept. Generally Cloud provides various functionalities like read write and copy of the original data. Owner of data first set the policies for the their data which data owner wants to place in cloud using attribute based encryption and send it to cloud service provider enclosed in JAR files. Any access to the data by user will be automatically checked for its authentication and logs record for each data item will be created by logger and sent to data owner for monitoring the data usage. If any unauthorized actions are performed by any user then immediately that action will be informed to owner by automatic reporting mechanism. It would generate the random numbers set for every user along with the data owner. So if the user accessing the account, the user has to give the random number set and that will be verified by the cloud server only. If the resultant verification is positive then and then only the users will be allowed to access their account. To reduce the overhead in proving access control based on user, there is need to use attribute based access control. Based on particular users attributes, access policy will be defined. User matching to this attributes will be given access to data items. So by using this way user management as well as user revocation will become easy.

Logger, Jar module, Attribute based encryption are the main components of accountability whereas loggers main tasks include automatically logging access to data item, it also control the data access even after it is downloaded by some user. Jar module generates single Jar for each data item, when the data item is uploaded. This release the overhead of generating multiple inner jar and ABE used for authentication. Auditing mechanism contains push and pull mode Whereas push mode refers to logs being periodically sent to the data owner pull mode refers to an alternative approach whereby the users can retrieve the logs as needed.

B. Comparison With Available Techniques:-

After doing detail study of already available techniques, we can found advantages as well as some limitations of those papers. In [1] authors only work on image data type not for other types. So in our propose idea we may applied this accountability concept to generic file types which contains all data types also it supports a variety of security policies like usage control for executables, indexing policies for text files. Their concept is shown in following fig.3:

They also use CIA concept, above figure define overview of their whole concept with steps. As per our observations there is some limitations in there flow like when we download data from cloud and when we use that data that time it saves log information in our file and it uploads that information only when network available but it is wrong because when we use that file offline that time owner never knows about this usage of his file. In our idea we give offline access of log information to owner. In paper they do not provide hashing of log files so in proposed idea this will overcome which is necessary for faster retrieval of logs.

In [2] as per our observation it has one limitation that users data which is processed on cloud server is in unencrypted format means not any encryption technique used for that data at the time of processing so there is a risk of data leakage. So it is very important to overcome that problem. In our proposed idea we used strong encryption techniques like attribute base encryption which provide strong security to data. Data will process on cloud in encrypted format only hence no need to worry about data leakage and also protect user's data from hackers.

In [3] describes how the privacy manager used to protect data. As per our observation, when input data is very large in size then this method can still require a large amount of memory. Also it does not provide guaranteed protection once the data are being disclosed. So this problem will overcome in propose idea. In [4] shows some leakage problems with the users data because it requires the server to send the linear combinations of data blocks to the auditor. In our idea this problem will overcome and it guaranties that there is no data leakage.

In [5] there is one limitation of Continuous auditing of agent. Proposed idea provides solution that monitors incorrect behavior of agent. Also in our proposed idea Log files should be reliable and proof to avoid illegal deletion, insertion and modification by third parties. Strong recovery mechanisms also use to restore damaged log files caused by some technical problems. Also as per our view the proposed technique should not intrusively monitor data recipients systems, nor should it introduce heavy communication overhead, which otherwise will hinder its feasibility.

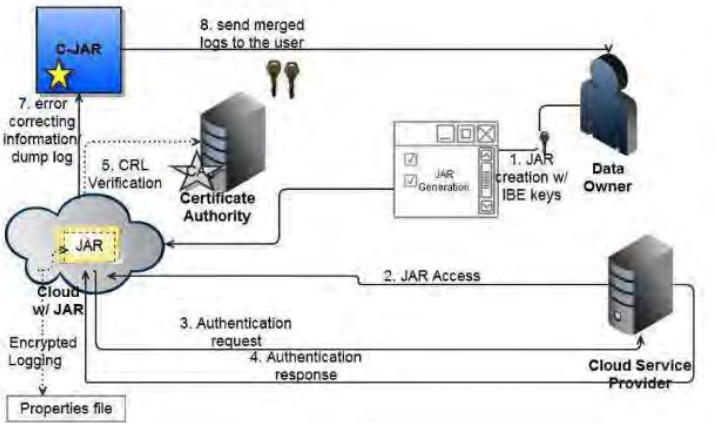

Figure.3: Existing system with CIA framework

IV. CONCLUSION

Cloud information accountability framework effectively monitors all the usage of the data in cloud. Hence by studying all related papers, we conclude that there is need to design the secure system using Cloud Information Accountability framework to meet the Accountability, Auditability and Authentication requirement for data sharing in the cloud environment which provides strong protection to cloud data and also overcomes the problems of previous techniques.

REFERENCES

[1] Smitha Sundareswaran, Anna C. Squicciarini and Dan Lin, "Ensuring Distributed Accountability for Data Sharing in the Cloud," IEEE Transaction on dependable a secure computing, VOL. 9, NO. 4, pg 556-568, 2012.

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

[3] S. Pearson, Y. Shen, and M. Mowbray," A privacy Manager for Cloud Computing," Proc. Int'l Conf. Cloud Computing (cloudcom), pp.90-106, 2009.

[4] Q. Wang, C. Wang, K. Ren, W. Lou and J. Li," Enabling public auditability and data dynamics for storages security in cloud computing", in INFOCOM.IEEE,2010,pp. 525-533.

[5] 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, pp. 187-201, 2005.

[6] Xiao Zhang, Hong-tao Du ,Jian-quan Chen, Yi Lin, Lei-jie Zeng "Ensure Data Security in Cloud Storage IEEE" 2011.

[7] Yubo Tan , Xinlei Wang "Research of Cloud computing Data Security Technology IEEE", 2012.

[8] P L Rini, Anand N " Encoding Personal Information On Data Sharing In Cloud Using BASE64 Algorithm" GRET 2013.

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

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

[11] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attributebased encryption," in Proc. IEEE Symp. Security and Privacy, Oakland, CA, 2007.

[12] Ryan K L Ko et al. "TrustCloud: A Framework for Accountability and Trust in Cloud computing," HPL-2011-38.