IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014
ISSN: 2320 – 8791 (Impact Factor: 1.479)
www.ijreat.org

A New Approach for Confidentiality Leakage Using DPPS in Cloud

Mrs.K.Kanimozhi[1], Mrs.S.Sankari[2]

[1]Final Year ME, Department of Computer Science and Engineering, S.Veerasamy Chettiar College of Engineering, Puliangudi, Tamilnadu, India

[2]Assistant professor, Department of Computer Science and Engineering, S.Veerasamy Chettiar College of Engineering, Puliangudi, Tamilnadu, India

Abstract

The main objective of this project is to provide encryption for privacy preservation. The users has large volume of intermediate datasets and by encrypting all intermediate datasets will lead to high overhead and low efficiency, when they are frequently accessed or processed of encryption and decryption. The users, before outsourcing, will first build an encrypted intermediate datasets, and then outsource the encrypted collection to the cloud server. So the users outsources the encrypted form only for selected intermediate datasets to the cloud server. To get back the encrypted intermediate datasets, user acquires a corresponding key. For user privacy logging plays a very important role in the proper operation of an organization's information processing system. However, maintaining logs securely over long periods of time is difficult and expensive in terms of the resources needed. So we proposed a complete system to securely outsource log records to a cloud provider by a new cloud computing paradigm, Data protection Privacy service (DPPS) is a suite of security primitives offered by a cloud platform, which enforces data security and offers evidence of privacy to data owners.

Index Terms— *cloud computing, intermediate datasets, privacy preservation.*

I. INTRODUCTION

Cloud Computing as a computing model, not a technology. In this model "customers" plug into the "cloud" to access IT resources which are priced and provided "on-demand". Essentially, IT resources are rented and shared among multiple tenants much as office space, apartments, or storage spaces are used by tenants. Delivered over an Internet connection, the "cloud" replaces the company data center or server providing the same service. Thus,

Cloud Computing is simply IT services sold and delivered overthe Internet.Cloud Computing vendors combine virtualization (one computer hosting several "virtual" servers), automated provisioning (servers have software installed automatically), and Internet connectivity technologies to provide the service. These are not new technologies but a new name applied to a collection of older (albeit updated) technologies that are packaged, sold and delivered in a new way. A key point to remember is that, at the most basic level, your data resides on someone else's server(s). This means that most concerns (and there are potentially hundreds) really come down to trust and control issues.

There are different types of cloud computing and are given as follows,

SaaS (Software As A Service):Is the most widely known and widely used form of cloud computing. It provides all the functions of a sophisticated traditional application to many customers and often thousands of users, but through a Web browser, not a "locally-installed" application. Little or no code is running on the Users local computer and the applications are usually tailored to fulfill specific functions.
SaaS eliminates customer worries about application servers, storage, application development and related, common concerns of IT. Highest-profile examples are Salesforce.com, Google's Gmail and Apps, instant messaging from AOL, Yahoo and Google, and VoIP from Vonage and Skype.

*PaaS (Platform as a Service) :*Delivers virtualized servers on which customers can run existing applications or develop new ones without having to worry about maintaining the operating systems, server hardware, load balancing or computing capacity. These vendors provide APIs or development platforms to create and run applications in the cloud – e.g. using the Internet.

Managed Service providers with application services provided to IT departments to monitor systems and downstream applications such as virus scanning for e-mail are frequently included in this category. Well known providers would include Microsoft's Azure, Salesforce's Force.com, Google Maps, ADP Payroll processing, and US Postal Service offerings.

IaaS (Infrastructure as a Service): Delivers utility computing capability, typically as raw virtual servers, on demand that customers configure and manage. Here Cloud Computing named as Randomized Efficient Distributed (RED) Protocol and Linear Hash Table (LHT) protocol. provides grids or clusters or virtualized servers, networks, storage and systems software, usually (but not always) in a multitenant architecture.

IaaS is designed to augment or replace the functions of an entire data center. This saves cost (time and expense) of capital equipment deployment but does not reduce cost of configuration, integration or management and these tasks must be performed remotely. endors would include Amazon.com (Elastic Compute Cloud [EC2] and Simple Storage), IBM and other traditional IT vendors

Valuable intermediate datasets need to be stored for sharing or reuse. It build an intermediate data dependency graph (IDG) from the data provenances in scientific workflows.With the IDG, deleted intermediate datasets can be regenerated, and as such we develop a novel algorithmthat can find a minimum cost storage strategy for the intermediate datasets in scientific cloud workflowsystems.

2. CLOUD ENVIRONMENTS

Cloud computing is a model for enabling service user's ubiquitous, convenient and on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services), that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing enables cloud services.

Nowadays, we have three types of cloud environments: Public, Private, and Hybrid clouds. A public cloud is standard model which providers make several resources, such as applications and storage, available to the public. Public cloud services may be free or not. In public clouds which they are running applications externally by large service providers and offers some benefits over private clouds. Private Cloud refers to internal services of a business that is not available for ordinary people. Essentially Private clouds are a marketing term for an architecture that provides hosted services to particular group of people behind a firewall. Hybrid cloud is an environment that a company provides and controls some resources internally and has some others for public use. Also there is combination of private and public clouds that called Hybrid cloud.

3. THREAT MODEL

The goal of our work is that an intruder who has complete access to the database server for some time should learn very little about the data stored in the database and the queries performed on the data

Our trust and attack model is as follows:
1. We do not fully trust the database server because it may be vulnerable to intrusion. Furthermore, we assume that, once a database intruder breaks into the database, he can observe not only the encrypted data in the database, but can also control the whole database system. A number of query messages sent by the user, as well as the database's processing of these queries, can be observed by the intruder. We note that assumption that an intruder can only control the whole database system for only a bounded time period is reasonable, for example, in the setting that a database administrator can physically reset the database server from time to time or when intrusions are detected.
2. We assume the communication channel between the user and the database is secure, as there exist standard protocols to secure it We also trust the user's front-end program; protecting the front-end program against intrusion is outside of the scope.
3. We require all data and metadata, including user logs and scheme metadata, to be stored encrypted. (Otherwise, these may open the door for intruders.)

4. RELATED WORK

The strategy achieves the best trade-off of computation cost and storage cost by automatically storing the most appropriate intermediate datasets in the cloud storage. [1]. With the IDG, deleted intermediate datasets can be regenerated, and as such we

develop a novel algorithm that can find a minimum cost storage strategy for the intermediate datasets in scientific cloud workflow systems. The strategy achieves the best trade-off of computation cost and storage cost by automatically storing the most appropriate intermediate datasets in the cloud storage. This strategy can be utilised on demand as a minimum cost benchmark for all other intermediate dataset storage strategies in the cloud. [2]. This approach alone may lead to excessive data distortion or insufficient protection. Privacy-preserving data publishing (PPDP) provides methods and tools for publishing useful information while preserving data privacy. Recently, PPDP has received considerable attention in research communities, and many approaches have been proposed for different data publishing scenarios[3]. solve the challenging problem of privacypreserving multi-keyword ranked search over encrypted cloud data (MRSE).We establish a set of strict privacy requirements for such a secure cloud data utilization system [7]. formulate and address the problem of authorized private keyword searches (APKS) on encrypted PHRin cloud computing environments. [10].

Encrypting all intermediate data sets will lead to high overhead and low efficiency when they are frequently accessed or processed. As such, we propose to encrypt part of intermediate data sets rather than all for reducing privacy-preserving cost. To propose a new cloud computing paradigm, Data protection Privacy service (DPPS) is a suite of security primitives offered by a cloud platform, which enforces data security and privacy and offers evidence of privacy to data owners, even in the presence of potentially compromised or malicious applications.Such as secure data using encryption, logging, and key management.The privacy to user will be protected in an secure manner.The privacy is interrupted by key management scheme.

5. PROPOSED SYSTEM

Data Protection Privacy Service improves secutity to users by enhanced logging mechanism.Securely maintaining log records over extended periods of time is very important..

Delegating log management to the cloud appears to be a viable cost saving measure. In this paper, we identify the challenges for a secure cloud-based log management service.

A new cloud computing paradigm, ***Data protection Privacy service (DPPS)*** is a suite of security primitives offered by a cloud platform is proposed, which enforces data security and privacy and offers evidence of privacy to data owners, even in the presence of potentially compromised or malicious applications. Such as secure data using encryption, logging, and key management.

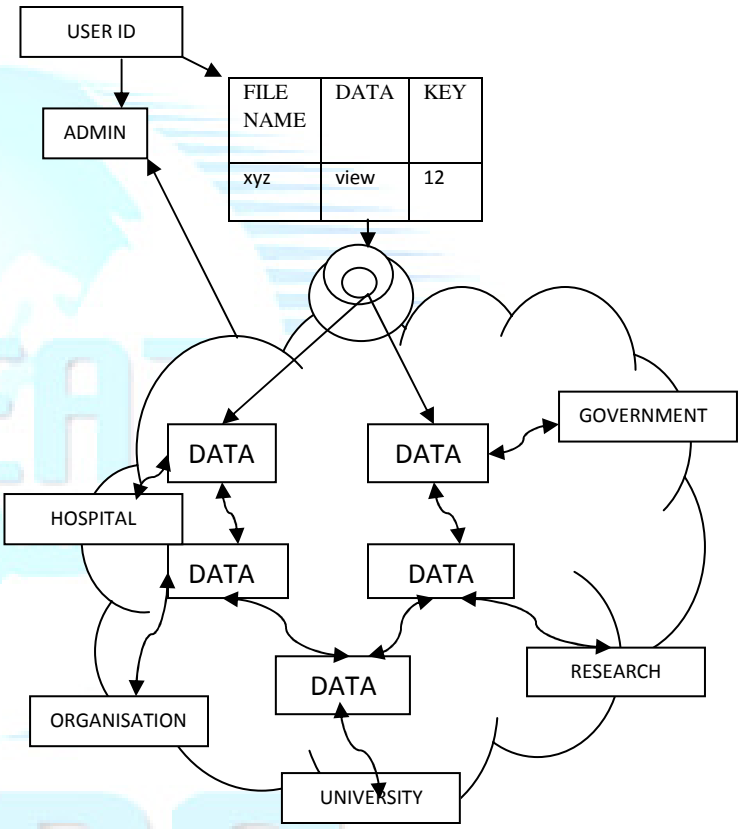

Figure.1 A general system model for our work

Considering a cloud data hosting service involving two different entities, as illustrated in Figure. 1. The users has intermediate datasets and outsourced to the cloud server in the encrypted form only for selected intermediate datasets.

The users, before outsourcing, will first build an encrypted intermediate datasets, and then outsource the encrypted collection to the cloud server. To getback the encrypted intermediate datasets , authorized user acquires a corresponding key.

We begin by summarizing the desirable properties that we seek from a secure log management service based on the cloud

computing paradigm. We will subsequently analyze our framework against these properties.

1) *Correctness*: Log data is useful only if it reflects true history of the system at the time of log generation. The stored log data should be correct, that is, it should be exactly the same as the one that was generated.

2) *Tamper Resistance*: A secure log must be tamper resistant in such a way that no one other than the creator of the log can introduce valid entries.

In addition, once those entries are created they cannot be manipulated without detection. No one can prevent an attacker who has compromised the logging system from altering what that system will put in future log entries.

One cannot also prevent an attacker from deleting any log entries that have not already been pushed to another system. The goal of a secure audit log in this case is to make sure that the attacker cannot alter existing log entries (i.e., the pre-compromise log entries) and that any attempts to delete or alter existing entire will be detected.

3) *Verifiability*: It must be possible to check that all entries in the log are present and have not been altered. Each entry must contain enough information to verify its authenticity independent of others.

If some entries are altered or deleted, the ability to individually verify the remaining entries (or blocks of entries) makes it possible to recover some useful information from the damaged log. Moreover, the individual entries must be linked together in a way that makes it possible to determine whether any entries are missing.

4) *Confidentiality*: Log records should not be casually browse able or searchable to gather sensitive information. Legitimate search access to users such as auditors or system administrators should be allowed.

In addition, since no one can prevent an attacker who has compromised the logging system from accessing sensitive information that the system will put in future log entries, the goal is to protect the pre-compromised log records from confidentiality breaches.

5) *Privacy*: Log records should not be casually traceable or linkable to their sources during transit and in storage.

There are three types of entities in our system:

• Users: Authorized users are able to read, write and search encrypted data residing on the remote server. Sometimes we may need to revoke an authorized user. After being revoked, the user is no longer able to access the data.

• Server: The main responsibility of the data storage server is to store and retrieve encrypted data according to authorized users' requests.

• Key management server (KMS): The KMS is a fully trusted server which is responsible for generating and revoking keys. It generates key sets for each authorized user and is also responsible for securely distributing generated key sets.

When a user is no longer trusted to access the data, the KMS revokes the user's permission by revoking his keys. Authorized users are fully trusted. They are given permissions to access the shared data stored on the remote server by the data owner. They are believed to behave properly and can protect their key sets properly.

- The initialization algorithm Init(1k) is run by the KMS which takes as input the security parameter 1k and outputs master public parameters Params and a master key set MSK.
- The user key sets generation algorithm Keygen (MSK,i) is run by the KMS which takes as input the master keys MSK and a user's identity i, generates two key sets Kui and Ksi. Kui is the user side key set for user i and Ksi is the server side key set for user i.
- The data encryption algorithm Enc(Kui,D,kw(D)) is run by a user who uses his key set Kui to encrypt a document D and a set of keywords associated kw(D), then outputs ciphertext c□i (D,kw(D)).
- The data decryption algorithm Dec(Kui,c′ i (D))is run by a user which decrypts c′ i (D) by using the user's key set and outputs D.

6. SECURITY ANALYSIS

We can prove the security of our scheme by using standard cryptographic techniques. Recall that for security we need to consider t queries. Suppose the uth query ($1 \leq u \leq t$) is of the format "select Aju,1 ,. . . , Aju,ℓ from T where Aju,0 = vu." We show that our basic solution only reveals j1,0, .. . , jt,0 beyond the minimum information revelation. That is, the only extra information leakage by the basic solution is which attributes are tested in the "where" conditions.

The security of our scheme derives from the security of the block cipher we use. In cryptography, secure block ciphers are modeled as pseudorandom. Here, encryption key of the block cipher is the random seed for the pseudorandom permutation.

7. CONCLUSION AND FUTURE WORK

The proposed system uses encryption and by encrypting all intermediate data sets will lead to high overhead and low efficiency when they are frequently accessed or processed of encryption and decryption. So we encrypt part of intermediate data sets rather than all for reducing privacy-preserving cost. Logging plays a very important role in the proper operation of an organization's information processing system. However ,maintaining logs securely over long periods of time is difficult and expensive in terms of the resources needed. So the privacy to data owners was given by Data protection Privacy service (DPPS). It is a suite of security primitives offered by a cloud platform, which enforces data security and privacy and offers evidence of privacy to data owners.

REFERENCES

[1] S.Y. Ko, I. Hoque, B. Cho, and I. Gupta, "Making Cloud Intermediate Data Fault-Tolerant," Proc. First ACM Symp. Cloud Computing (SoCC '10), pp. 181-192, 2010.

[2] D. Yuan, Y. Yang, X. Liu, and J. Chen, "On-Demand Minimum Cost Benchmarking for Intermediate Data Set Storage in Scientific Cloud Workflow Systems," J. Parallel Distributed Computing, vol. 71, no. 2, pp. 316-332, 2011.

[3] Data Publishing: A Survey of Recent Developments," ACM Computing Survey, vol. 42, no. 4, pp. 1-53, 2010.

[4] H. Takabi, J.B.D. Joshi, and G. Ahn, "Security and Privacy Challenges in Cloud Computing Environments," IEEE Security & Privacy, vol. 8, no. 6, pp. 24-31, Nov./Dec. 2010.

[5] D. Zissis and D. Lekkas, "Addressing Cloud Computing Security Issues," Future Generation Computer Systems, vol. 28, no. 3, pp. 583- 592, 2011.

[6] H. Lin and W. Tzeng, "A Secure Erasure Code-Based Cloud

[7] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM '11, pp. 829-837, 2011.

[8] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," Proc. 41st Ann. ACM Symp. Theory of Computing (STOC '09), pp. 169-178, 2009.

[9] M. Li, S. Yu, N. Cao, and W. Lou, "Authorized Private Keyword Search over Encrypted Data in Cloud Computing," Proc. 31st Int'l Conf. Distributed Computing Systems (ICDCS '11), pp. 383-392, 2011.

[10] B.C.M. Fung, K. Wang, R. Chen, and P.S. Yu, "Privacy-Preserving Data Publishing: A Survey of Recent Developments," ACM Computing Survey, vol. 42, no. 4, pp. 1-53, 2010.

[11] X. Zhang, C. Liu, J. Chen, and W. Dou, "An Upper-Bound Control Approach for Cost-Effective Privacy Protection of Intermediate Data Set Storage in Cloud," Proc. Ninth IEEE Int'l Conf. Dependable, Autonomic and Secure Computing (DASC '11), pp. 518-525, 2011.

[12] K.P.N. Puttaswamy, C. Kruegel, and B.Y. Zhao, "Silverline: Toward Data Confidentiality in Storage-Intensive Cloud Applications," Proc. Second ACM Symp. Cloud Computing (SoCC '11), 2011.

[13] K. Zhang, X. Zhou, Y. Chen, X. Wang, and Y. Ruan, "Sedic: Privacy-Aware Data Intensive Computing on Hybrid Clouds," Proc. 18th ACM Conf. Computer and Comm. Security (CCS '11), pp. 515-526, 2011.

[14] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," Comm. ACM, vol. 53, no. 4, pp. 50-58, 2010.