Effficient Audit Service Outsourcing For Data Integrity In Clouds

Syeda Imrana Fatima
ME,CSE I-Year
Everest Educational Society's Group Of Institutions
College Of Engineering & Technology,
Computer Science & Engineering
Department,Aurangabad.

Seema singh solanki
Professor,
Everest Educational Society's Group Of Institutions
College Of Engineering & Technology,
Computer Science & Engineering
Department,Aurangabad.

Abstract:

Cloud computing is an emerging scenario in today's world. With the advent of new technologies new challenges associated with them to emerge, as in cloud computing. Cloud computing is faced by challenging issues like data security, data integrity, data duplication, authentication and authorization. Providing data integrity is a tricky task in cloud computing.

the intricacy faced by the user at the time of storage management and storage maintenance can be reduced by cloud based data outsourcing in which a hassle free platform for data storage is provided which is of considerable low-cost, is scalable and is location-independent. To guarantee data integrity audit service are vital, audit services plays a significant role to ensure the integrity and accessibility of outsourced data and to achieve digital forensics and reliability on cloud computing.

In order to verify the integrity of a given text or document Provable data possession (PDP), a cryptographic technique can be used. This can be done without retrieving the data at an un-trusted server, this technique can be used to carry out audit services. We intend to build a interactive PDP protocol to prevent duplicity, deceitfulness , treachery , untruthfulness, deceit and dishonesty and corrupt data.

We intent to came up with a efficient mechanism with respect to probabilistic queries and periodic verification to reduce the audit costs per verification and implement abnormal detection timely. We also intend to create amethodology by which wecan select an optimal threshold value to reduce computational overheads of cloud audit services. This ensures high availability and integrity of data, along with content security and user privacy.

Keywords:

cloud computing, data integrity in cloud Security, Cloud storage, Interactive proof system Provable data possession Audit service streaming, RTP,

1. Introduction:

With the new advent in technology, the traditional information systems are getting an easy substitute in the form of cold computing. Cloud computing is a rapidly growing and fast evolving technique. Cloud computing provides a scalable environment for budding amounts of data and that work is carried out on various applications and services by means of on-demand self-services. One of the fundamental characteristic of this paradigm shifting is that data are being centralized and outsourced into clouds.

The outsourced storage service provides a comparably low-cost, scalable, location-independent platform for managing clients' data. Hence storage management and storage maintenance are taken care of by cloud storage service (CSS). Whereas clouds are quite susceptible to crashes or attacks or failures which could be irreparable, irretrievable, irreversible. It could incur huge loss of important and useful data.

The main reasons for these risks is that the cloud infrastructures are much more powerful and reliable than personal computing devices However, they are still vulnerable to security threats both from outside and inside the cloud (Armbrust et al., 2010); for the benefits of their possession, there exist various motivations for cloud service providers (CSP) to behave unfaithfully toward the cloud users (Tchifilionova, 2011); furthermore, the dispute occasionally suffers from the lack of trust on CSP.

Consequently, their behaviors may not be known by the cloud users, even if this dispute may result from the users' own improper operations (Ko et al., 2011). Therefore, it is necessary for cloud service providers to offer an efficient audit service to check the integrity and availability of the stored data (Yavuz and Ning, 2009). Traditional cryptographic technologies for data integrity and availability, based on hash functions and signature schemes (Hsiaoet al., 2009; Yumerefendi and Chase, 2007), cannot work on the outsourced data without a local copy of data. More so ever, it is not a convenient solution for data validation.

As might require downloading them which might be expensive especially for large-size files. Moreover, the solutions to audit the correctness of the data in a cloud environment can be formidable and expensive for the cloud users (Armbrust et al., 2010). Therefore, it is critical to realize public audit ability for CSS, so that data owners may resort to a third party auditor (TPA), who has expertise and capabilities that a common user does not have, for periodically auditing the outsourced data. This audit service is significantly important for digital forensics and data assurance in clouds(Yan Zhu 2012).

2. CLOUD COMPUTING: overview, issues, and challenges:

Cloud computing one of the latest promising technology trends today, for its potential to be a "unsettling" technology. The goal of this section is to address the unique challenges and data integrity threats of cloud computing for practical application and utilization of Cloud Computing.

The data storage and computing are not in the local computer and server but in the amount of computer distributed in the internet in the cloud computing. The cloud computing move the tasks which are implemented in the personal computer and private data center into the larger computing center which are shared with total user and distributed in the internet.

It compose applications out of loosely coupled services and one service failure will not disrupt other services. The cloud computing system can be divided into two sections: the front end and the back end. They connect to each other through the internet. The front end is user who use the service provided by the back end which is the cloud section of the system. The cloud is a metaphor for the Internet, based on how it is depicted in computer network diagrams, and is an abstraction for the complex infrastructure it conceals.

If the environment is built correctly, virtual servers will not be affected by the loss of a host. Hosts may be removed and introduced almost at will to accommodate maintenance. The virtual servers in the cloud computing system can be scaled out easily and if the administrators check out that the resources supporting a virtual server are being taxed too much in the real environment and they can modify the amount of resources allocated to that virtual server. The user need not computing and storage resource and don't provide the application in the cloud computing. The resource and server can be provided by the cloud computing.

The cloud computing can be classified into into private cloud, public cloud and hybrid cloud based upon the difference of service object. The hybrid cloud is the composition of two or more clouds and bounded by standard or proprietary technology. Hybrid clouds combine character of both public and private clouds. The private cloud is deployed in the company and the security can be made easily.

Private clouds are virtualized cloud data centers inside firewall and it is a private space dedicated to system within a cloud data center. Private cloud refers to internal data centers of a business or other organization not made available to the general public. The cloud system infrastructures are owned by an organization which sells cloud services to the general public or to a large industry company. The public cloud is running in the internet and the security is very complex. Public clouds are virtualized data centers outside of firewall and the service provider makes resources available to consumer on demand over the public Internet. The cloud computing is highly virtualized and standardized infrastructures and it can give more efficient and application management.

It has the character of massive scalability and it can deliver more applications to large number of users. Cloud computing allows for flexibility, and capital and operational expenses for resources are only incurred when they are needed. The cloud computing is on-demand service and it give computing capabilities as needed automatically. It can use the service by many machine such as desktop, laptop, PDA and mobile phone.

The cloud service model include SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service).In the software as a service the consumer use the provided application and don't manage or control the network, servers storage and the application.

It can reduce expenses and is easy to use and access everywhere. It share instance of a software application as a service accessible via internet browser or client based role access and sharing rules. The service provider hosts the software so the user don't need to install or manage or buy hardware for it. All they have to do is connect and use it.

The examples of SaaS are Flickr, Google Docs, Siri, Amazon and Cloud Drive. In platform as a service the consumer deploys their applications on the cloud computing system and controls their applications but they don't manage servers and storage and delivers a computing platform or solution stack as a service.

It share platform for custom software application configuration, development, testing and deployment. The examples of PaaS are Google App Engine, Amazon Web services. In the infrastructure as a service the consumer get access to the infrastructure to deploy their application and system but they don't manage or control the infrastructure and they control the storage and applications. It share managed pool of configurable and scalable resources such as network, middleware, database and storage servers. The examples of IaaS is Amazon Elastic Compute Cloud (EC2).

The cloud has the elastic character and resource allocation can get bigger or smaller depending on demand. The cloud also has the scalability and the cloud can scale upward for peak demand and downward for lighter demand There are many cloud computing systems in the market such as Google, Windows, IBM and Amazon. The Google cloud computing systems include GFS (Google File System), Map Reduce and Bitg table.

3. CLOUD INTEGRITY PROBLEM:

As the cloud system runs over the internet security issues faced by the internet can also be faced by the cloud system. The cloud systems are quite similar to traditional systems i.e. pc and are vulnerable to special and new security issues. The major concerns about cloud computing are data integrity and privacy. The traditional security problems such as security vulnerabilities, virus and hack attack can also make threats to the cloud system and can lead more serious results because of property of cloud computing.

Hackers and malicious intruder may hack into cloud accounts and steal sensitive data stored in cloud systems. The data and business application are stored in the cloud center and the cloud system must protect the resource carefully. Cloud computing is a technology evolution of the widespread adoption of service oriented architecture, virtualization and utility computing over the Internet and it includes the applications, platform and services. If the systems meet the failure, fast recovery of the resource also is a problem. The cloud systems hide the details of service implementation technology and the management.

The user can't control the progress of deal with the data and the user can't make sure the data security by themselves. Data moving to any authorized place you need it, in a form that any authorized application can use it, by any authorized user, on any authorized device. Data integrity requires that only authorized users can change the data and Confidentiality means that only authorized users can read data.

Cloud computing should provide strong user access control to strengthen the licensing, certification, quarantine and other aspects of data management. In the cloud computing, the cloud provider system has many users in a dynamic response to changing service needs. The users do not know what position the data and do not know which servers are processing the data. The user do not know what network are transmitting the data because the flexibility and scalability of cloud system.

The user can't make sure data privacy operated by the cloud in a confidential way. The cloud system can deploy the cloud center in different area and the data can be stored in different cloud node. The different area has different law so the security management can meet the law risk. Cloud computing service must be improved in legal protection.

4. Proposed System:

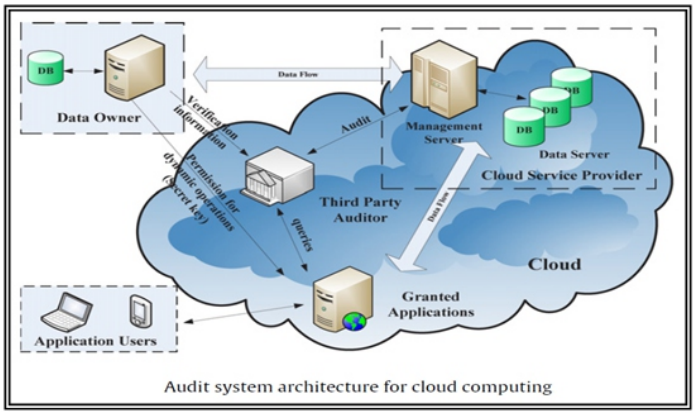
To come up securely and steadily an efficient third party auditor (TPA), the following two essential requirements have to be met:

1) The TPA should without introducing any further online difficulty to the cloud user be able efficiently audit the cloud data storage without demanding the local copy of data.

2) The auditing process carried third party should not bring in new vulnerabilities towards user data privacy.

PDP (public Provable data possession), which is a cryptographic technique is utilized for verifying the integrity of data without retrieving it at an un- trusted server; can be used to realize audit services. a random mask technique to achieve a privacy-preserving public auditing system for cloud data storage protection while keeping all above requirements in mind.

Architecture:



Audit system architecture for cloud computing

We use bilinear aggregate signature in order to carry out efficiently multiple auditing tasks, as well as to extend our main result into a multi-user site, where TPA can perform multiple auditing tasks concurrently. Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient. We can also extent our main scheme to support batch auditing for TPA upon delegations from multi-users.

Modules:

1.Audit Service System.
2.Data Storage Service System.
3.Audit Outsourcing Service System.
4.Data integrity and Performance Analysis.

Audit Service System:

A robust cryptographic interactive audit scheme ensures public audit which has to be carried out periodically. It is essential to ensure that we provide cryptographic interactive such a way so as to ensure that the audit well retains the accuracy property and zero-knowledge property of proof systems. These properties ensure that our scheme can not only prevent fraud and falsification of cloud storage data of different providers, but also prevent the seepage of outsourced data in the process of verification.

Data Storage Service System:

In this module, we considered FOUR entities to store the data in secure manner:

1.Data owner (DO)
This is basically the client who has a large amount of data to be stored securely over the cloud.

2.Cloud service provider (CSP)
The cloud service provider ensures protected data storage service and has enough storage spaces and computation resources.

3.Third party auditor (TPA)
Third party auditor has capability to administer or scrutinize – outsourced data under the entrustment of data owner.

4.Granted applications (GA)
Who after the generation of security key has the right to access and manipulate stored data. These applications can be either inside clouds or outside clouds according to the specific requirements.

Audit Outsourcing Service System:

In this module the data owner after the generation secret key uses it to preprocess the file, which might consists of a compilation of blocks, generates a set of public verification information that is stored in TPA, transmits the file and some verification tags to Cloud service provider CSP, and may delete its local copy.

At a later time, using a procedure of verification of irretrievability, TPA (as an audit agent of clients) issues a test to audit (or check) the integrity and availability of the outsourced data in terms of the public verification information. It is necessary to give an scrutinize the data for abnormal events.

Data integrity and Performance Analysis:

In this we try to find out the data integrity and performance analysis of our proposed system.

• Audit-without-downloading

To allow TPA (or other clients with the help of TPA) to verify the correctness of cloud data on demand without retrieving a copy of whole data or introducing additional on-line burden to the cloud users.

• Verification-correctness

Cloud Service provider shoud not be able to pass no cheating CSP that can audit from TPA without storing users' data intact.

• Privacy-preserving

To ensure that there exists no way for TPA to derive users' data from the information gathered during the auditing process.

• High-performance

To allow TPA to execute auditing with least amount of overheads in storage, communication and computation, and to support statistical audit sampling and optimized audit program with a long enough period of time.

5. CONCLUSIONS:

On the one hand Cloud Computing offers some implausible benefits like unlimited storage, scalability, elasticity, platform independent, low-cost and reliability. access to quick processing power and the ability to easily share and process information, on the other hand though, it does have several issues, and most of which are security related.

Cloud systems must overcome many short comings before in order to be widely accepted. Several security issues currently affect cloud systems, however, there may be many undetermined, unstipulated, unspecified and undiscovered security issues. Therefore there is still a need for optimal solutions if cloud systems are to be widely adopted. Main problems that need to be addressed are regarding data encryption, data privacy , data integrity these problems hinder the development of cloud computing and the security issue is the core problem.

In this paper, discussed the construction of an efficient audit service for data integrity in clouds. We proposed an interactive audit protocol to implement the audit service based on a third party auditor.in this the TPA issues a periodic verification to scrutinize outsourced data . for this the security of TPA has to be maintained This approach greatly reduces the workload on the storage servers, while still achieves the detection of servers' misbehavior with a high probability

References:

Yan Zhua,b, Hongxin Huc, Gail-Joon Ahnc, Stephen S. Yauc 2011 Efficient audit service outsourcing for data integrity in clouds Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson.

D.A., Rabkin, A., Stoica, I., Zaharia, M., 2010. A view of cloud computing. Commun. ACM 53 (4), 50–58.

Ateniese, G., Burns, R.C., Curtmola, R., Herring, J., Kissner, L., Peterson, Z.N.J., Song, D.X., 2007. Provable data possession at untrusted stores. In: Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, pp. 598–609.

Ateniese, G., Pietro, R.D., Mancini, L.V., Tsudik, G., 2008. Scalable and efficient provable data possession. In: Proceedings of the 4th International Conference on Security and Privacy in Communication Networks, SecureComm, pp. 1–10.

Barreto, P.S.L.M., Galbraith, S.D., O'Eigeartaigh, C., Scott, M., 2007. Efficient pairing computation on supersingular abelian varieties. Des. Codes Cryptogr. 42 (3), 239–271.

Beuchat, J.-L., Brisebarre, N., Detrey, J., Okamoto, E., 2007. Arithmetic operators for pairing-based cryptography. In: Cryptographic Hardware and Embedded Systems– CHES 2007, 9th International Workshop, pp. 239–255.

Boneh, D., Boyen, X., Shacham, H.,2004. Short group signatures. In: In Proceedings of CRYPTO 04, LNCS Series. Springer-Verlag, pp. 41–55.

Boneh, D., Franklin, M., 2001. Identity-based encryption from the weil pairing. In: Advances in Cryptology (CRYPTO'2001). Vol. 2139 of LNCS, pp. 213–229.

Bowers, K.D., Juels, A., Oprea, A., 2009. Hail: a high-availability and integrity layer for cloud storage. In: ACM Conference on Computer and Communications Security,pp. 187–198.

Cramer, R., Damgård, I., MacKenzie, P.D., 2000. Efficient zero-knowledge proofs of knowledge without intractability assumptions. In: Public Key Cryptography, pp. 354–373.

Dodis, Y., Vadhan, S.P., Wichs, D., 2009. Proofs of retrievability via hardness amplification. In: Reingold, O. (Ed.), Theory of Cryptography, 6th Theory of Cryptography Conference, TCC 2009. Vol. 5444 of Lecture Notes in Computer Science.Springer, pp. 109–127.

Erway, C.C., Küpc¸ ü, A., Papamanthou, C., Tamassia, R., 2009. Dynamic provable data possession. In: Proceedings of the 2009 ACM Conference on Computer and Communications Security, CCS 2009, pp. 213–222.

Fu, K., Kaashoek, M.F., Mazières, D., 2002. Fast and secure distributed read-only file system. ACM Trans. Comput. Syst. 20 (1), 1–24. Goldreich, O., 2001. Foundations of Cryptography: Basic Tools. Vol. Basic Tools. Cambridge University Press.

Hsiao, H.-C., Lin, Y.-H., Studer, A., Studer, C., Wang, K.-H., Kikuchi, H., Perrig, A., Sun, H.-M., Yang, B.-Y., 2009. A study of user-friendly hash comparison schemes. In: ACSAC, pp. 105–114.

Hu, H., Hu, L., Feng, D., 2007. On a class of pseudorandom sequences from elliptic curves over finite fields. IEEE Trans. Inform. Theory 53 (7), 2598–26**05.**