Cloud Computing and Protected Storage Services Technique

Neha Chaudhary
M.Tech. Scholar, Department of Computer Science & Engineering
Krishna Institute of Engineering & Technology

Vineet Sharma
Associate Professor, Department of Computer Science & Engineering
Krishna Institute of Engineering & Technology

ABSTRACT

Cloud Computing is the Shared pool of resources, which can be requested on-demand to enjoy the services and application. The data owner remotely store their data services and application which can be retrieved when required. The clients save their data on the storage services and gets free from the burden of maintaining the data. But some of the incidences in the history of cloud computing shows that they are not reliable at times Ex: Gmail disaster, Amazon S3 downfall. Some of the major challenging issues of the cloud computing are availability of service, data lock in, Data confidentiality and Auditability, Data transfer Bottleneck etc.

The data outsourcing makes the owner free from data storage, physical organization and security, which is the need of both enterprises and individuals with high service level requirements. In order to smooth the progress of rapid deployment of cloud data storage service and regain security assurances with outsourced data dependability, efficient methods that enable on-demand data correctness verification on behalf of cloud data owners have to be designed. While public audit ability, we can have a trusted third audit party who have knowledge and capability to evaluate the danger of outsourced data whenever demanded. This is a cost effective method for data owners to believe on the cloud. We have challenges that need to be resolved for publicly auditable secure cloud storage services to become a reality.

Keywords— MAC, PDP, POR, TPA.

1. INTRODUCTION

The cloud computing refers to the services and application delivered through internet. The software and hardware are embedded in the data centres providing those services. The data centres are also called cloud (comprise of hardware and software). The Nantional Institute of Standard and Research has given the standard definition of Cloud Computing which is being Accepted Worldwide:-

"Cloud Computing is a model for enabling convenient , 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 services provider interactions. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models".[1]

Essential characters tics of cloud computing comprises of five points (1) on-demand self service (2) Broad network access (3) Resource pooling (4) Rapid elasticity (5) Measured services. Along with Saas cloud computing also provides Paas (platform as a service) and Iaas (Infrastructure as a service). Cloud can be deployed in various forms like: private cloud (owned by an organization and data centres are not available to general public), public cloud (data centres are available to general public and the organization sell the services in pay-as-you-go manner), community cloud (shared by a no. of organizations and provide services to a specific community), hybrid cloud (composition of one or more clouds)[1].

The cloud computing gives the following illusion of the hardware (1)Infinite computing resources available on demand.(2) Elimination of an up-front commitment by cloud users, therefore allows users to start and increase hardware resources on demand. (3) Use of computing resources on short term basis on demand as needed therefore conserving resources. The cloud is deployed in various manner and the services provided to the public is utility computing. Examples of utility computing include Amazon web services, Google application engine, and Microsoft ozure. The services of cloud computing is provided as Saas (software as services). Saas proves beneficial to both the end user and services providers. It generalize the procedure of software installation to the services provider and the end user can access the services "anytime, anywhere". The deployment of Saas and scale on demand without building and provisioning a data centre [2].

One of the greatest challenge of the cloud computing is verification of integrity. The POR (proof of retriveability) provides back-up services to produce the proof that a user is retrieving a reliable data. POR can be referred as a cryptographic proof of knowledge. The main aim of POR is to provide checks on verification without having to download the whole file. Verification of the authenticity is also one of the major concerns on un-trusted servers. The facility of PDP (Provable data possession) allows a client to have trust on the servers without retrieving it. PDP generates probabilistic proofs of possession on sets of blocks from the servers. Thus protocol transfer small amount of data, thus reducing the network communication. Thus PDP provides remote data checking support. But practically implementing PDP protocols is limited by disk I/O not by cryptographic computation [5]. In the current scenario of Saas the customer cannot take of risk of losing data stored by relying on these services. The third party auditing is important to evaluate the risk and to increase the efficiency of insurance based risk migration. Various approaches are designed to support both internal and external auditing of online storage services. The service provider and the auditors have to adopt these approaches.

2. RELATED WORK

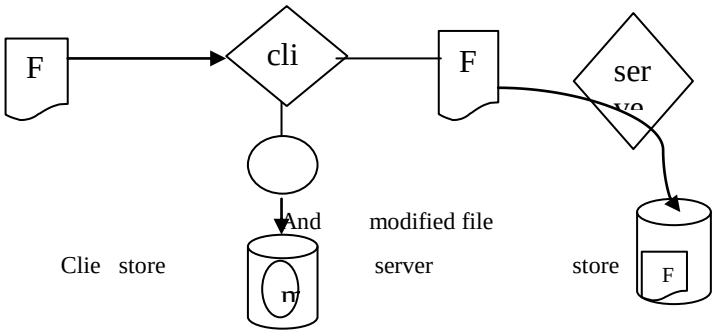
The research of cloud computing is to create a storage system which combine the advantage of scaling resources utility up and down as demanded as well as meeting programmers expectations in regard to resource management for scalability, data durability and high availability. The cloud computing provides the facility of "pay-as-you-go" which proves

economic to the buyers. Elasticity helps to reduce the wastage of resources and compensate for high cost servers how paying as you go is economic than buying. With the various advantages of cloud computing there are obstacles which prevent adoption and growth of cloud computing. The challenging issue in cloud computing is the verification of the integrity of the data retrieved from the cloud. Various models are proposed by the researchers and the models are able to achieve the public auditability efficiently without retrieving the file to the TPA. The various models are (1)Basic Model using MAC (2)Provable Data Possesion(POR). (3)Proof of Retrieveability(POR).

The Basic Model of security Using MAC: We start from the very basic solution that could form the basis of public auditing services for dependable cloud data storage. One of the basic approach to protect the data integrity would be using traditional cryptographic method, the well-known message authentication codes (MACs). Initially, data owners can locally keep a small amount of MACs for the data files to be outsourced. Whenever the data owner needs to get back the file, it can verify the integrity by recalculating the MAC of the received data file and comparing it to the locally pre-computed value. While this method allows data owners to verify the correctness of the received data from the cloud, it does not give any assurance about the accuracy of other outsourced data. The scheme does not give any guarantee that the data in the cloud are all undamaged, unless the data are all downloaded by the owner. Because of the large cloud data, it would be quite impossible for a data owner to get back all of her data for verification.

The PDP Model: This model allows the client to verify its data integrity without retrieving it. Considering that the file data are large and are stored at re- mote sites, accessing an entire file is expensive in I/O costs. Reading an entire archive, even periodically, greatly limits the scalability of network stores. Furthermore, I/O incurs to establish data possession interferes with on-demand bandwidth to store and retrieve data. We conclude that clients need to be able to verify that a server has retained file data without retrieving the data from the server and without having the server access the entire file. The model of PDP provides probabilistic proof that a third party stores a file. The homomorphic property, tags computed for multiple file blocks can be combined into a single value. The client is going to pre computes the tags for each block of a file and then stores the file and its tags on server. When the client wants to verify that the server possesses the file it can generate a random challenge against a randomly selected set of file blocks. Using the queried blocks and their corresponding tags, the server generates a proof of possession. The client is thus convinced of data possession, without actually retrieving file blocks.

Input file client generates metadata(m)



(a)Pre-processing and store

(1) Client generates a random challenge

(2) Server computes proof of possession

(3) Security verifies secure proof.

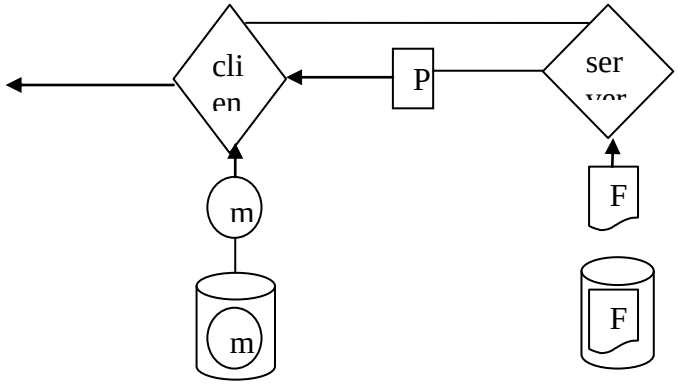


Fig 1:Protocol for Provable data possession

The PDP Scheme provides probabilistic proof that a third party store file. It provides the data format independence that put no limitation on the no. of challenges being executed.

The POR Model: The exploration of POR(proof of retrievability) a cryptographic building block to user through which user get a protocol interface which help in retrieving data in its entirety. POR protocol stores only a single cryptographic key—irrespective of the size and number of the files whose retrievability it seeks to verify—as well as a small amount of dynamic state (some tens of bits) for each file. POR protocol encrypts F and randomly embeds a set of randomly-valued check blocks called sentinels. POR protocol encrypts F and randomly embeds a set of randomly-valued check blocks called sentinels. The use of encryption here renders the sentinels indistinguishable from other file blocks. The verifier challenges the prover by specifying the positions of a collection of sentinels and asking the prover to return the associated sentinel values. If the prover has modified or deleted a substantial portion of F, then with high probability it will also have suppressed a number of sentinels. It is therefore unlikely to respond correctly to the verifier.

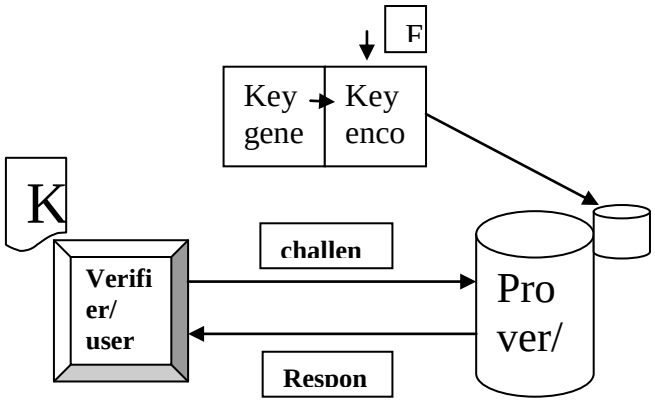


Fig 2: Working of POR system

We let ˜ F refer to the full, encoded file stored with the prover. A drawback of our proposed POR scheme is the pre-processing / encoding of F required prior to storage with the prover. This step imposes some computational overhead—beyond that of simple encryption or hashing—as well as larger storage requirements on the prover. The sentinels may constitute a small fraction of the encoded ˜ F (typically, say, 2%); the error-coding imposes the bulk of the storage overhead. For large files and practical protocol parameterizations, however, the associated expansion factor | ˜ F|/|F| can be fairly modest, e.g., 15% [5].

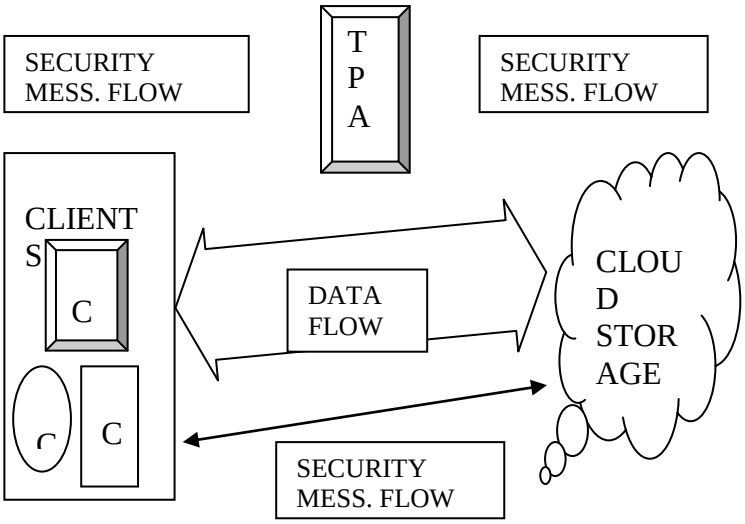
Security model with the help of Third Party Auditiour(TPA): One of the solution of data's verification is to assign the work to a trusted third party and then the data owner is free from the work of maintenance and computing of the security of data. The TPA is a special entity with extra capability to verify the integrity of the data and extract the data from cloud to client with having a local copy of the data. The previous model studied allows the TPA to do the verification of the data. But they are not feasible if the data dynamics are performed in the storage ex: Insertion, deletion and block modification. The security model leads to security loopholes.

The Cloud data storage Architecture:-

Client: The client is the one which store its data on the cloud and relies on the cloud for data maintenance and computation. The client can represent either the individual user or the organization.

Cloud Storage Service: The Cloud Storage Service are provided by the CSP (Cloud service provider). The storage services store the clients data and are responsible for the maintenance and the computation of the data.

Third Party Auditor(TPA): The TPA is the one with special capability to audit the cloud for verification of data. The client assign the task of verification to TPA and gets free from the extra computation.



Fig[3]: cloud storage service architecture

The TPA thus help in verifying the data on the cloud on-behalf of the data owner. The computation of the dada owner is thus reduces and the owner gets free from the responsibility of verifying the data. But the drawback of all the models is that they are not feasible with data dynamic operations. The implementation of the data dynamics in the POR and PDP lead to security loopholes.

3. CONCLUSION

Cloud Computing represents the opportunity for more income at little extra cost. Although Cloud Computing providers may run afoul of the obstacles, but over the long run it will successfully navigate these challenges and set an example for others to follow, perhaps by successfully exploiting the opportunities that correspond to those obstacles. When file blocks are MACed, it is effectively possible to convert an erasure code into an error correcting code. The decoding process simply discards corrupted blocks. Motivating the need for auditing to support an online service-oriented economy, highlight the issues around both internal and external auditing and detail ways of auditing online storage services. PDP is used to prove possession of large amounts of data. Experiments show that such schemes also impose a significant I/O and computational burden on the servers. By utilizing the homomorphic token with distributed verification of erasure coded data, protocol achieves the integration of storage correctness insurance and data error localization, i.e., whenever data corruption has been detected during the storage correctness verification across the distributed servers, we can almost guarantee the simultaneous identification of the misbehaving servers. The crypto system of the POR model help us to verify that the prover(cloud) store the file and the Verifier can retrieve the file upon after the successful; completion of the protocol as the prover posses the considerable propobality of correct file data. With the help of sentinel the verifier check the possibility of corrupted data and based on the probability returned from the prover it can judge for the integrity.One of the basic drawback of this system is that it is also applicable for static data and not dynamic data integrity checking. If the verifier would like to modify a few data blocks then the archive would delete or modify the set of blocks with impunity, having learned that they are not sentinels.

4. FUTURE PROSPECTS

The evolution of cloud computing leads to revolution in IT market. It has opened the doors to utilize any type of service anywhere and anytime. Although cloud computing proves to be cost efficient as well as reduces lot of maintenance power to the clients. Besides the benefits of cloud computing it also prone to many drawbacks and among those one of the crucial one is the security of data. Various security models are designed to secure the authenticity and for verification of data but some of them proves to have security loopholes. The Security Models can be enhanced to provide the Data Dynamic operations which help in verifying the dynamic data. The Models can also be extended to provide the Batch auditing so that the computation cost of the TPA can be reduces as it support the Multi client environment. All the models being developed consider the TPA to be trustworthy but they can also cheat the owner. So, the Models can also be enhanced by considering the TPA to be dishonest.

5. REFERENCES

[1] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing," 2009; http://csrc.nist.gov/groups/SNS/cloud-computing/index.html

[2] M. Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," Univ. California, Berkeley, Tech. Rep. UCBEECS-2009-28, Feb. 2009.

[3] Janakiram MSV "Demystifying the Cloud, introduction to Cloud Computing" Version 1.1, 2010.

[4] A. Juels, J. Burton, and S. Kaliski, "PORs: Proofs of Retrievability for Large Files," Proc. ACM CCS '07, Oct. 2007, pp. 584–97.

[5] G.Ateniese et al., "Provable Data Possession at Untrusted Stores," Proc. ACM

[6] Mehul A. Shah, Mary Baker, Jeffrey C. Mogul, Ram Swaminathan, "Auditing To Keep Online Services Honest"; http://www.hpl.hp.com/personal/Mehul_Shah/papers/hotos11_2007_shah.pdf.

[7] Cong Wang, Qian Wang, and Kui Ren, "Ensuring data storage security in cloud computing", Dept of ECE Illinois Institute of Technology; http://ece.wpi.edu/~wjlou/publication/IWQoS09_Wang.pdf

[8] Amazon.com, "Amazon s3 Availability Event: July 20, 2008," July 2008; http://status.aws.amazon.com/s3-20080720.html

[9] M. Arrington, "Gmail Disaster: Reports of Mass Email Deletions," Dec. 2006; http://www.techcrunch.com/2006/12/28/gmail-disaster-reports-of-massemail- deletions/

[10] G. Ateniese et al., "Scalable and Efficient Provable Data Possession," Proc. SecureComm '08, Sept. 2008

[11] C.Wang et al.,"Ensuring Data Storage Security in Cloud Computing," Proc. IWQoS '09, July 2009, pp. 1–9.

[12] William stallings,2007 "*Cryptography and Network Security, Principles and Practices*",Volume, Prentis Hall.