Dynamic Data and Mutual Trust For Cloud Computing Storage System

Vaishali B. Kolate
M.E. student Computer Engg
PVPIT ,Pune, India.
Vaishali.bkolate@gmail.com

Abstract : **Many organizations are producing a large amount of data which they need to store But storing such a large data is somewhat problematic at a local storage area. So this system provide storage as a service facility and cloud service provider. This system allow the different organizations to store data at remote site by paying some amount as per the usage. Thus it minimizes the overhead of storing a large data at local storage.**

Keywords-Storage as a service , cloud service provider.

I: INTRODUCTION

Cloud computing has important advantages including: cost effectiveness, low management overhead, immediate access to a wide range of applications, flexibility to scale up and down information technology (IT) capacity, and mobility where customers can access information wherever they are, rather than having to remain at their desks. It is a distributed model over a large pool of shared-virtualized computing resources. Cloud service providers offer different classes of services Storage-as-a-Service , Application-as-a-Service, and Platform-as-a-Service that allow organizations to concentrate on their core business .

Currently different organizations produces different information like personal , electronic health data, financial transactional information. Digital data amount is also increasing so rapidly . This data needs to be get distributed over a wide area as local management of such huge amount of data is problematic and costly. Storage-as a service offered by this cloud service provider is used to avoid maintenance cost of different business and provide high storage facility. CSP provide this storage facility for customers in exchange of fees measured in GB/month. Because of this storage system different owners store their data on remote server instead of local storage area . CSP provide the recovery system for stored data for this facility it store different duplicate copies of data on different sites. Because of this storage facility different authorized users can access their data remotely from any location. As owners store their sensitive data to CSP storage they want confidentiality, integrity, and access control of their data. Data confidentiality is very important issue. For example, in e-Health applications the data should have privacy and it should follow some policies so that is should not display any personal information to unauthorized users.

II : LITERATURE SURVEY

In distributed networks different techniques are available like integrity, cryptography, and access control for the data. PDP protocol is implemented for sensitive data. Different PDP schemes are available for different data which is stored dynamically at different locations. Examples of PDP schemes that deal with dynamic data are .This scheme is only for one copy of data but as we have implemented PDP we can use it for multiple copies too. In current area storage facility is provided to local storage area but as it is very complicated and costly too .We need to implement such storage facility which provide vast storage area for different data provided by the different owners.

Aameek Singh describe “SHAROES” it is a platform for data sharing in the storage-as-a-service model. SHAROES uses novel cryptographic access control primitives (CAPs) to support rich data sharing semantics without trusting the SSP for enforcement of security policies. He showed how SHAROES is able to support an expressive access control model, which in conjunction with its in-band key management technology provides seamless transition ability from local storage to the outsourced model with minimal user involvement. [1]

Giuseppe Ateniese introduced a model for provable data possession, in which it is desirable to minimize the file block accesses, the computation on the server, and the client-server communication. They incur a low (or even constant) overhead at the server and require a small, constant amount of communication per challenge.[2]

Francesc Sebe provide first practical protocol for remote file integrity checking allowing an infinite number of verifications presented. An ordering or a structure between the set of files should be defined, so that the set

of files can be regarded as a super file. Once the super file is defined, its integrity can be checked using his protocol without any modification.[3]

E. Goh present SiRiUS, which is designed to be layered over existing file systems such as NFS to provide end-to-end security. To enforce access control in SiRiUS, each data file is attached with a metadata file that contains an encrypted key block for each authorized user with some access rights . More specifically, the md-file represents the d-file's access control list . The d-file is encrypted using a file encryption key, and each entry in the ACL contains an encrypted version of the FEK under the public key of one authorized user.[8]

Zhuo Hao has proposed new remote data integrity checking protocol for cloud storage. The proposed protocol is suitable for providing integrity protection of customer's important data. The proposed protocol supports data insertion, modification and deletion at the block level, and also supports public verifiability. [4]

Khaba M.V explored the problem of providing simultaneous public audit ability and data dynamics for remote data integrity check in Cloud Computing. His construction is deliberately designed to meet these two important goals while efficiency being kept closely in mind.[7]

Venkata Pallavi presented an efficient PDP scheme for distributed cloud storage. Based on homomorphism verifiable response and hash Index hierarchy, a cooperative PDP scheme to support dynamic scalability on multiple storage servers is proposed. Her scheme provides all security properties required by zero knowledge interactive proof system, so that it can resist various attacks even if it is deployed as a public audit service in clouds.[5]

Vaishnavi V R proposed a model which detects the probable parameters for detecting blocks of distributed files that reduce the overheads across the evaluation of veracity verification structure of queries which are seen as major obstacle in health care Systems. The probable frequency over protocol verification obtained from block distribution for estimating the actual values. The main purpose is attained by reducing the costs for overhead according to its computation with dynamic verification system. It enhances the security over performance oriented data storage that is outsourced to external servers.[6]

III: PROPOSED CLOUD BASED STORAGE SCHEME

Cloud based storage system contain different component :

Data owner : this is nothing but an organization which generate sensitive data which is getting stored on cloud storage.

Cloud servers :Different cloud servers and also provide facility to store different data on different servers.

Trusted third party : It is one which is accepted by all the participants and this follows different policies for detecting an unauthorized party.

Authorized user : This are the different users which have permission to access the data from different owners.

Example: In medical sensitive data can be the different patient information like date of admission, decease detected ,doctor recommended , prescribed medicines and so much information. In this medical system the different doctors are the users who have authority to access data , medical center is the owner of data who provide the data, and different organizations are nothing but the trusted third parties

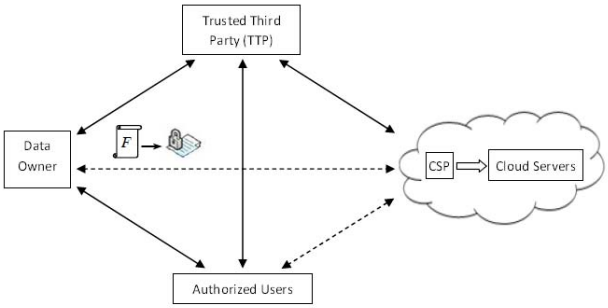

Fig 1: The Cloud computing data storage system .

If we outsourced the data to the unauthorized CSP then it is very difficult to access it back .Because of this data owners have issue of integrity of stored data. It may happen that a dishonest owner may claim that the data stored in the cloud is corrupted to get some compensation from CSP . If mutual trust between data owner and CSP is not maintained then it will affect the CSP deployment. We can use the digital signature to detect unauthorized users. When any owner stored their file on the remote site it attaches its tag with that file which we call owner tag. The owner tags are generated per block not per file to enable dynamic operations at the block level without retrieving the whole outsourced file. The owner sends its file to the CSP, where the tags are first verified. If verification failed , the CSP can not store the data blocks and asks the owner to re-send the correct tags. If the tags are valid, both the blocks and the tags are stored on the cloud servers. The tags achieve non-repudiation of the file . When any user want to access the file it request it from the CSP in response of that request CSP send the signature and tag for that file . The authorized user first verifies the tags . If verification failed the CSP need to perform the process verification process. After receiving the file user check its tag. After submitting file back CSP checks its tag if verification is not successful then it shows that file is get corrupted. The CSP cannot reflect such corruption in data blocks.

CSP is using the digital signature for detecting the dishonest users. CSP also offer the assurance of newness property. As large amount of data has been stored on the remote site so to identify the data it is attached with the owner tag. As computation overhead is increased on different system components the data owner generates a signature for each block. For each outsourced data the CSP perform the digital signature verification. For each received block the authorized users also perform the signature verification at the cloud servers. For a file F which is containing m blocks, the solution need 2m signature generations and 3m signature verifications, which will be computationally a challenging task for large data files. Take an example of the outsourced file which is of size 1GB with 4KB block size. For solution requires 219 signature generations and 3 × 218 signature verifications. If the CSP receives the data blocks from a trusted entity the block tags and the signature operations are not needed . As we are delegating a small part of the owner's work to the TTP this reduces storage and computation overheads. The outsourced data should be kept private and should avoid any possible leakage of data .

Overview of system :

Our scheme in this addresses important issues related to outsourcing data storage. This issues are data dynamic, newness, mutual trust, and access control. The data owner is only allowed to update the outsourced data file. For validating dynamic data ,its newness property requires the knowledge of some metadata that reflects the most recent modifications issued by the owner. Different indices are used to reflect that the file has been added, updated, deleted, and inserted at the requested position. In CSP we combine hash values and data structure this is nothing but block status table. The TTP is used to provide the mutual trust between different system components. In this proposed scheme we use three cryptographic techniques: bENC, lazy revocation, and key rotation. The bENC enables a data owner to encrypt some secret information to only authorized users allowing them to access the outsourced data file. Lazy revocation is used to revoked users which can read unmodified data blocks, when updated/new blocks are encrypted under new keys generated from the secret information broadcast to the authorized users. key rotation is used by the authorized users because of which they can able to access both updated/new blocks and also unmodified ones that are encrypted using older key than current key.

System setup diagram:

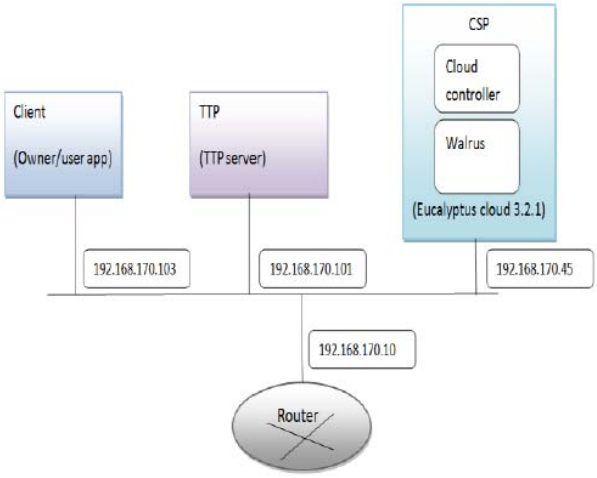

Fig 2. cloud based storage system diagram

Notations:

F : is a data file to be outsourced, and is composed of a sequence of n blocks, i.e., F = {b1 b2 , . . . , bn }

h : is a cryptographic hash function

DEK : is a data encryption key

EDEK :is a symmetric encryption algorithm under DEK, e.g., AES (advanced encryption standard)

F : is an encrypted version of the file blocks

F HTTP :is a combined hash value for F , and is computed and stored by the TTP

T HTTP: is a combined hash value for the BST, and is computed and stored by the TTP

Ctr: is a counter kept by the data owner to indicate the version of the most recent key

Rot = ctr, bENC(Kctr) is a rotator, where bENC(Kctr) is a broadcast encryption of the key Kctr

⊕: is an XOR operator Procedural Steps of the Proposed Scheme Setup and File Preparation.

The system setup is done for lifetime of the system and it is done from both side that is owner side and TTP side.

Owner Role:

Data Owner can able to store the data at remote site. They have ability to add , delete, update and insert some data on remote file. In the initial step the owner initialize the ctr counter to 1 and add the file . It generate the signature and tag for file and attach it with the file which is to be get added . He can rotate the the counter forward and backsword . because of this owner can get the file from storage and also can upload the updated file back to storage. Owner encrypt the file and then upload that file on the remote site. Owner can delete the file from storage.

TTP :

TTP is working as a intermediate between owner and the CSP . when file is getting stored in the CSP it first get send to the TTP . TTP compute the hash values of the file and also hash value of the TTP . TTP attaches this values with the file and then send this file to the CSP.

TTP provide the assurance of the newness property and intergrity.If any user want to access the data from remote site even if owner is offline , still this system provide this facility.

Operations performed on the data

Modification:

Data owner can able to modify the file which is stored on the remote site. When any file owner want to modify it is calculating the hash value for the new file and combine the old value with the new value. After modifying the file if replaces the old hash value with the new value.

Append:

We can append the data block with the file which is stored. Append operation is done at the end of the last block of the file.

Deletion:

We can delete the file block from the storage. This is a very easy process. In this when any file get deleted from the storage its hash value is also get deleted, and all the indices are get changed and moved upwards.

Data Access and Cheating Detection:

In this operation the user who want to access the file send request to the CSP . In response of this request the CSP generate two signature for the file. Only authorized users can access the file . If file tag not match and file signature is not verified then cheating is also get detected.

IV: MATHEMATICAL MODEL

Data owner :

1.Increament counter ctr

2.Generate initial secret key

3.Generate the encrypted file version of the file

4.Generate the tag for file

5.Create BST

6.It generate the rotator which can be used by the revoked users to get old version of the file .Owner send the encrypted file ,BST , and rotator to the TTP.

V: EXPECTED RESULT

If we perform the experimental result of the proposed system then expected result will be as follows

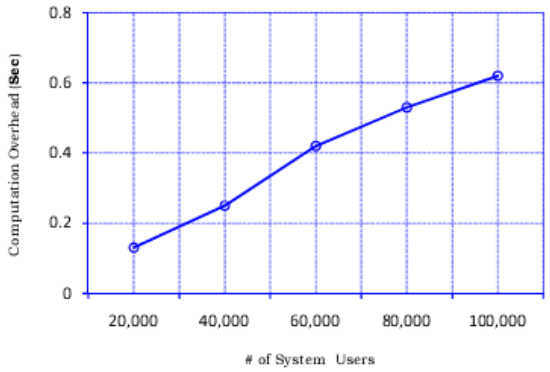

Fig 3:owner's average computation overhead due to dynamic operation

The computation overhead of the system of storing data and accessing the data will get decreased. And increased the response time and access time. The above diagram shows the expected result curve of the system . it shows that the at less system overhead it serves more users .

VI : SECURITY PROVIDED BY SYSTEM

In our proposed system there are some security measures which are satisfied

Data confidentiality:

Only authorized users can access the file from CSP. As owner is using the signature and tag are attached with the file it provide the confidentiality feature of data.

Assurance of newness property.

This property is very important in different operations like insertion, deletion operations. This is nothing but the system should reflect the updated version of the file when user is accessing the file and at the same time someone else is doing some modifications in it.

Enforcement of access control.

In this system we combine lazy revocation, key rotation, and broadcast encryption to provide access control of outsourced data. Revoked users are only allowed to access the unmodified block . If anyone make change in the file block then revoked users can not get access to the updated content.

Detection of dishonest owner/user

This is very important security feature which is satisfied by our system . in this CSP is generating a digital signature which used while communication in the system. So when any user does not satisfy the signature verification then it detect as dishonest user or owner.

VII : CONCLUSION

In this we proposed the cloud based storage system which provide the facility to store the large amount of data from different owners dynamically on different remote sites. This also provide facility to perform different operations on the file data block like modify , add, delete files. In this we have implements different cryptographic technique like lazy revocation, broadcast encryption, digital signature for providing the different security features to the data block.

ACKNOWLEDGEMENT

I have worked with with full support of my guide Prof. Y. B. Gurav . great spirit, very patiently with full interest on making high performance computing based project and eventually after a long research throughout, have made a project report entitled "Dynamic data and Mutual Trust for Cloud Computing Storage Systems"

REFERENCES

[1] A. Singh and L. Liu, "Sharoes: A data sharing platform for outsourced enterprise storage environments," in Proceedings of the 24th International Conference on Data Engineering, ICDE. IEEE, 2008.
[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song. Provable data possession at untrusted stores. Cryptology ePrint archive, May 2007.
[3] F. Sebe, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient remote data possession checking in critical information infrastructures," Knowledge and Data Engineering, IEEE Transactions on, vol. 20, pp. 1034 –1038, aug. 2008.
[4] Z. Hao and N. Yu, "A multiple-replica remote data possession checking protocol with public verifiability," in Data, Privacy, and E-Commerce, September 2010.
[5] Venkata Pallavi , E.Padma "Authentication based remote data possession in multi-cloud storage" International Journal of Science, Technology & Management Volume No 04, Special Issue No. 01, March 2015.
[6] Vaishnavi V R , Senduru Srinivasulu , Divya C "Data veracity verification and cryptography algorithm for health care systems using cloud technologies" International Journal Of Pharmacy & Technology Vol. 8 , Issue No.1 March-2016.
[7] Khaba M.V , M.Santhanalakshmi "Remote Data Integrity Checking in Cloud Computing" International Journal on Recent and Innovation Trends in Computing and Communication ISSN 2321 – 8169 Volume: 1 Issue: 6 553 – 557 JUNE 2013.
[8] E. Goh, H. Shacham, N. Modadugu, and D. Boneh, "SiRiUS: securing remote untrusted storage," in NDSS, 2003.
[9] F. Brick, "Are you ready to outsource your storage?" *Computer Technology Review*, June 2003.
[10] Amazon Storage Service., *http://aws.amazon.com/s3*.