www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 4 Issue 1 January 2015, Page No. 10144-10249

Improving Cloud Data Storage Using Data Partition and Recovery

Prof:-A.R.Zade, Shaikh Umar, Potghan Rahul, Rale Sagar and Borade Sagar
JSPM's Rajarshi Shahu College Of Enggineering Tathawade Pune-33
Department of Information Technology, Savitribai phule Pune University
umarlinc@gmail.com, rahulpotghan5727@gmail.com, sagarrale56@gmail.com, sagarborade96@gmail.com

Abstract- Cloud storage system allows storing of data in to the cloud server efficiently and makes the user to work with the data without any problem, trouble of the resources. Also the Cloud storage permits users to remotely store their data and enjoy it on _demand high quality cloud applications without the burden of local hardware and software management. In existing system data are stored in the cloud using dynamic data operation with computation which makes the user need to make a copy of data for further updating and verification of the data loss unfortunately.

In proposed System Storing the data on to cloud is partition, and Digital Signature (D.S) to each partition data and it stored into different server.

Keywords: - Remote Data Integrity Checking,
Error Localization, Partitioning, dynamic data, Cloud Storage.

I. INTRODUCTION

In today's high speed network, the Internet access becomes more popular and available in the recent years. Cloud computing is an totally internet based technology, which is used widely nowadays to enable the end user to create and use software without worrying about the execution of the technical information from anywhere as well as any time .

Storing data in a third party's cloud system causes serious disquiet over data privacy. General encryption systems protect data confidentiality, as well as limit the functionality of the storage system because a few operations are supported over encrypted data.

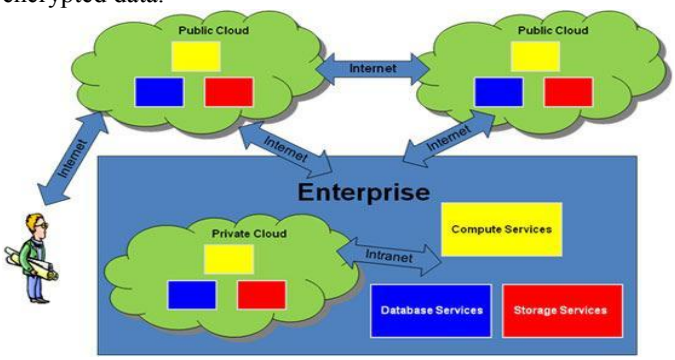

Fig1:- cloud services Architecture

Cloud storage is a service for designers to store and access the data in cloud. Data are stored in the cloud through hosted network services and also it offers the use of access control on it.

Types of Service Models in Cloud:-

Cloud computing providers offer their services according to three fundamental models. Infrastructure as a service (IaaS), software as a service (SaaS) where IaaS, and Platform as a Service (PaaS), is the most basic and each higher model abstracts from the details of the lower models.

1. *Software as a Service (SaaS):-*

Capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. Various client devices access the applications through client interface, like web browser, or a program interface. The consumer cannot manage or control the underlying cloud infrastructure including network, servers, os, storage, or even individual application capabilities, with the possible immunity of limited user-specific application configuration setting.

2. *Platform as a Service (PaaS): -*

Capability provided to the consumer is to deploy onto the cloud infrastructure consumer created or acquired applications created using, tools supported by the providers, programming languages. Consumer cannot manage or control the underlying cloud infrastructure including network, servers or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

3. *Infrastructure as a Service (IaaS):-*

Capability provided to the consumer is to provision processing, storage and networks, and also other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can contain OS and applications. Consumer cannot control the underlying cloud infrastructure but has control over OS, storage, and deployed applications; and possibly limited control of select networking components such as host firewalls.

Deployment Models in C C:-

Mainly four types of cloud existing in cloud com puting.

These deployment models describe who co ntrols and is responsible for the services. The detail types of different type cloud are as follows.

1. *Private cloud:-*

Cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consum ers. It may be controlled and operated by the organization, a third party, and it may exist ON or OFF premises.

2. *Public cloud:-*

Cloud infrastructure is provisioned for open use by the general public. It may be controlled and operated by a b usiness, or the government organizations. It exists on the pr emises of the cloud provider.

3. *Community cloud:-*

Cloud infrastructure is provisioned for exclusive use by a specific community of consumers from orgaanizations that have shared concerns. It may be controlled an d operated by one or more of the organizations in the com munity, a third party, and it may exist ON and OFF premises.

4. *Hybrid cloud:*

Hybrid cloud infrastructure is a composition of two or more different cloud infrastructures i.e. communi ty, public or private, that will be unique entities, but vault together by standardized technology that enables data a nd application portability (e.g., cloud bursting for load bala ncing between clouds).

Remote archive service is responsible for prop erly preserving the data; remote data integrity checking proto col detects the data corruption, losses and misbehaving server in the cloud storage. In the proposed work, data integrity checking is analyzed in internal and external ways. It supports data dynamics and public verification as in Fig [2] c onsidering the entrusted server with security analysis.

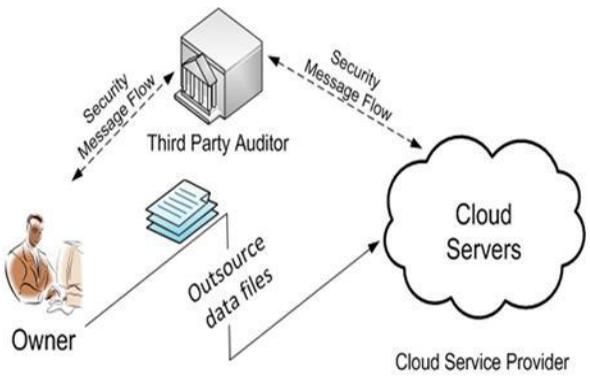

Fig:-2 data storage on to cloud

1. *Client(owner):-*

This entity, which has large data files to be store d in the cloud. Relies on the cloud for data maintenance and computation, can be individual, end user, consumers or organizati ons.

2. *Cloud Storage Server (C SS):-*

This entity, which is managed by Cloud Service Provider (CSP), has significant storage sp ace and computation resource to maintain the client's data.

3. *Third Party Auditor(TP A):-*

This entity, which has capabiliti es that clients do not have, is trusted to measure and uncover risk of cloud storage services on behalf of the clients upon requ est.

The advantage of the cloud stor age is flexible with reduced cost, space and they also manage the data loss risk and so on. This research works aims in d esigning to efficient flexible storage scheme to ensure the a vailability of data and data correctness in cloud, by partitio ning algorithm. Partitioning happens by usingRC6 algorithm. Storage and retrieval process are simplified by reducing the storage space when there is need to store and retrieved by merging technique

II. PROBLEM STATEMENT

A. *Design goal*

To ensure the data security and data storage efficiency in cloud, integrity checking is desig ned effectively.
Some phrases are being used in t he content that is as follows.

1. *Dependability*: - Improve the mechanisms work of the integrity checking a gainst the service attacks and threads.
2. *Lightweight*: - Communication and computation cost in sharing data and storage of the data in cloud is reduce.
3. *Error localization of data*: - Compute and consists fast access of the data an d detect the error.
4. *Storage*: - End user can store the data in to cloud at anytime and anywhere t hrough internet and easily.

B. *System model*

The different network Entities an d resources of Cloud storage service architecture with is repres ented as given below

1. *User*: - Enable e nd user to storing the data without any difficulty in to cloud.
2. *Cloud Server (CS)*:-It can Manage and provide storage space, preserve computational resources and storage serv ices by the cloud service provider (CSP).
3. *Remote data inte grity checking*: - Integrity checking to detect and correct the data error if loss it can be rec over and data localization in cloud data storage.

C. *Notations*

1. F -Data file in equal size stored in block wise.
2. En- Encoding the files a nd each consists n blocks.

3. Ind- Each separate block consists an index to represent the block when access.
4. FS - Data files are partitioned into portions and stored.
5. Pek - Generating the public key for encoding the files.
6. Pdk - Generating the private key to decode the files for access.
7. De - Decoding the files and consists the blocks. auditing of whether his data stored in the cloud are definitely

Table 1.

Notations	Descriptions
N	*Total number of data units*
k	*Minimum number of data units required for data retrieval*
p	*Total number of available service providers*
q	*Minimum numbers of service providers required for data retrieval*
i	*I=1,2,…,p*
SP_i	*Cloud service provider*
QS_i	*Quality of service factor for each service provider*
QS_{net}	*The QoS achieved at the time of retrieval*
C_i	*Cost of sorting per unit data for i^{th} service provider*
C_{tot}	*Total cost of storing the distributed customer data on p service provider*
a_i	*Number of data units that assigned to i^{th} service provider*
j	*J=1,2,.., a_i*
$x_{i,j}$	*j^{th} data unit on i^{th} service provider*

III. RELATED WORK

We did the literature survey for slicing mechanisms or data integrity checking and data storage if loss data that are currently used in dynamic multi transactional applications. [4] Dynamic data storage with token pre-computation and how it is stored in to cloud is analyzed which provide information about effective storage mechanisms. [4]

Cloud is dynamic, like electronic documents, or log files, etc. Therefore, it is essential to consider the dynamic case, where a user may hope to perform various block-level operations of update, delete, and modify the data file while maintaining the storage correctness assurance.

Many approaches are used to provide security to cloud .A plain approach is message authentication codes (MAC) can be used to protect the data integrity. Data owners will initially locally maintain a small amount of MAC for the data files which are to be outsourced. The data holder can verify the integrity by recalculating the MAC of the received data file when he or she wants to [4].

A hash tree can be working for large data files, in which leave contains hashes of data blocks and internal contains hashes of their children of the tree. So to perform this thing for data owner TPA can be used [3].

The work studies the problem of ensuring the integrity of data storage in Cloud Computing; we are considering the task of allowing third party auditor (TPA), on the cloud client, to authenticate the integrity of the dynamic data stored in the cloud.

TPA can be removes the involvement of the client through the auditing of whether his data stored in the cloud are definitely intact, which can be important in achieving economies of scale for Cloud Computing.

Integrity checking concepts is used to detect and recover, avoid miss behavior of server considering data correction and error localization [2].

Dynamic data process and public audit ability are used for supporting the data integrity, the objective of this work is to have independent perspective and quality in services evaluating with the third party auditor (TPA).

Storage model is also planned here to support multiple reviewing tasks to improve the efficiency, performance, security.

Data partitioning in vertical and horizontal directions as discussed in [1, 5]. They partitioned data into buckets and used slicing technique for data storage onto cloud. In the works [2], [4], developer considers generating signature methods for ensuring the cloud data storage security.

RC6 encryption of data and Token pre computation scheme ensures dynamic data operation and the integrity checking. This scheme provides data storage security.

The limitation with the existing mechanism it takes more time and cost to perform the dynamic processing of data encryption and decryption techniques to store data in cloud with the security.

IV. PARTITIONING ,INTEGRITY CHECKING FOR DATA STORAGE

The end users stores data in cloud and also they maintain data its own locally in the cloud data storage

1. *Cloud Storage*:-

Fig [3] shows how end user is supported with dynamic data operation and security model for storing data in to cloud. Unauthorized access is avoided it detects the threats and misbehaving of server and also prevents the data from attacks. propose of data storage architecture ensures pre-computation to check the corrections of the data This is done before storing the data and the dynamic data operation is done after the computation. This process enhances the security of data because the data are stored after the pre-computation process. By using pre-computation the security key is generated by the encryption technique to ensure security from unauthorized access key is generated by encryption and decryption technique to ensure security.

2. *Access(Retrieve) data from cloud storage*:-

Data are retrieved from the storage service as per the end user request or demand. As bellow Fig [3] the data can be retrieved or restored from the server ensuring the data correction. Decryption processes the key to reload the original data from the cloud server. The encoded data can be decoded for view the original data without pairing the re-encryption scheme.

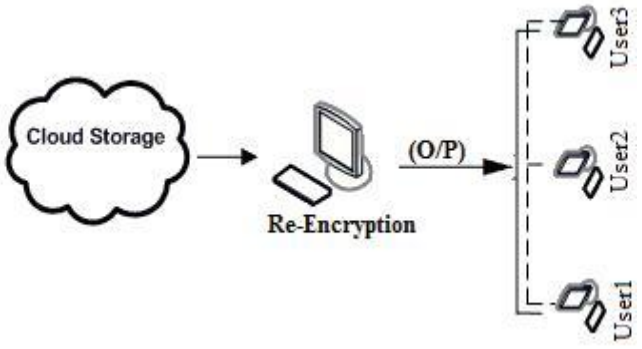

Fig:-3 data retrieve from cloud

The end user can also decide what data need to be accessed and shared by the other users in cloud. Data accessed from cloud service enables the services in secured manner.

During the retrieval of the data can be decrypted by generating the key and merging the data into the original data.

It also manages the effective storage and retrieval processes. Dynamic data operation, like as inserting, updating as well as deleting is also done before partitioning the data

3. *Partitioning data:-*

Partitioning function plays an important role in this work because it splits (break up) larger files into smaller parts to store the data effectively in quick manner enhancing easy access to data also when there is needed or demanded by end user. The original data is complex and there is difficulty in storing it in cloud, so partition function is used for make the storage easy in cloud. The partitioned files are encrypted, that is encoded with the key and stored in cloud. Partitioning done automatically when the data is fed for storing in cloud. Original file is also reassembled when there is need to access the same.

V. IMPLEMENTATION TOPICS

The important parameters like cost, storage time, and space and access control are also considered. Data are partitioned to storing it in cloud and combined during retrieval. Encryption and Decryption process are implemented to ensure security of the data. The concepts that have been implemented to achieve this goal have been discussed below.

1. *Public Audit ability*:-

This module provides the logical methods for handling attacks. It also supervises Error localization and misbehaving of server during the process of storing data securely. The Remote data
Integrity checking and secured data handling is done here.

2. *File Access:-*

The data are access from cloud as and when user needs or demand partitioned data is put together to view the original data before access during retrieval. After merging, the original data is decrypted with key. For each user key is generated to access the data from cloud.

3. *User Revocation:-*

The information of the key for accessing the data is conserved. When again encrypt the original data, the key and decryption server is generated for user and the information are secured and
Managed.

4. *Encryption:-*

Encryption technique is used to encrypt the files for the security. The file will be in cipher, by encrypting the file. A common approach is used to encrypt with RC6 (symmetric) key algorithm, SPEKE algorithm for key exchange

Algorithms:-

Encryption with RC6-w/r/b
Input:
Plaintext stored in 4 w-bit input registers A; B; C; D
Number r of rounds
W-bit round keys S [0; 2r + 3]
Output:
Cipher text stored in A; B; C;
D Procedure:
B = B + S[0]
D = D + S[1]
for i = 1 to r
do f
t = (B _ (2B + 1))<<<lg w
u = (D _ (2D + 1))<<<lg w
A = ((A _ t)<<<u) + S[2i]
C = ((C _ u)<<<t) + S[2i+ 1]
(A;B;C;D) = (B;C;D;A)
}
A = A + S[2r + 2]
C = C + S[2r + 3]E.

Decryption with RC6-
w/r/b Input:
Ciphertext stored in 4 w-bit input registers A;B;C;D
Number r of rounds
W-bit round keys S[0; : : : ; 2r +
3] Output:
Plaintext stored in A;B;C;D
Procedure: C = C -S[2r + 3]
A = A -S[2r + 2]
fori = r downto 1 do
f
(A;B;C;D) = (D; A;B;C)
u = (D _ (2D + 1))<<<lg w t
= (B _ (2B + 1))<<<lg w C
= ((C -S[2i + 1])>>>t) _ u
A = ((A -S[2i])>>>u) _ t
g
D = D -S[1]
B = B -S[0]

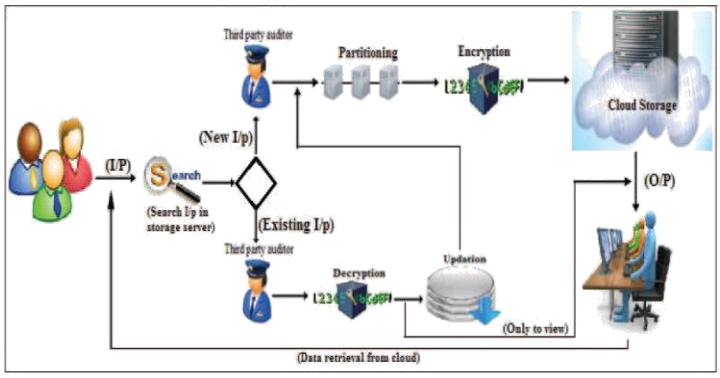

Fig:-4System overview

Performance Study

InFig. [5], the performance of the reduced space during storage of the partitioned data with data security is shown. Space reduction by 1.5kb, the time reduction during the data storage. By this way the data can be stored in a quick manner and the retrieval can effectively.

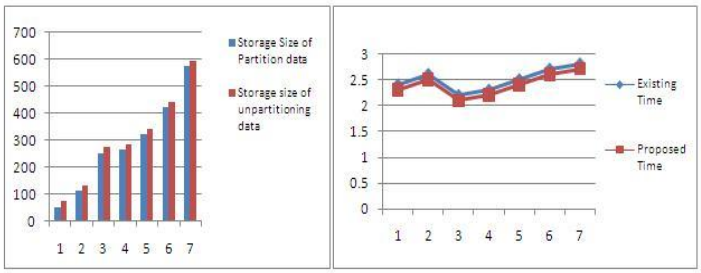

Fig 5 reduce time and space

VI. CONCLUSION AND FUTURE WORK

In this work utilizes the technique of separating the data in an efficient way, also provides high storage capacity along with less reduction of time. We used encryption and decryption techniques in our system which basically is an implementation. The Encryption and decryptions are basic steps used to maintain the security of the data. The prime benefit of our system includes the safety and protection of data along with maintaining the integrity of the data preserved.

We propose an efficient data storage security in cloud service. The partition process enables storing of the data in easily and effective manner. It also provides way for flexible access of data and there is less cost in data storage. The space and time is also effectually reduced during storage. Dynamic operation is another concept where, encoding and decoding process secures data, when storing into cloud. Also the remote data integrity checking detects the threats and misbehaving server while storing the data in cloud ensuring data security. Future work is planned for provide higher level of security and searching mechanisms for outsource computations in cloud services.

REFERENCES

[1] PDDS - Improving cloud data storage security using data partitioning technique Selvakumar, C. ; Rathanam, G.J. ; Sumalatha, M.R. Advance Computing Conference (IACC), 2013 IEEE 3rd
International
DOI: 10.1109/IAdCC.2013.6506806
Publication Year: 2013,

[2] Toward Secure and Dependable Storage Services in Cloud Computing Cong Wang ; Qian Wang ; Kui Ren ; Ning Cao ; Wenjing LouServices Computing, IEEE Transactions on Volume:5DOI:10.1109/TSC.2011.24
PublicationYear:2012,Page(s):220-232

[3] Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing Qian Wang ; Cong Wang;KuiRen;WenjingLou;JinLiParallel and Distributed Systems,IEEETransactionson
Volume:22,Issue:5
DOI:10.1109/TPDS.2010.183
Publication Year: 2011 , Page(s): 847 - 859

[4] Slicing: A New Approach for Privacy Preserving Data Publishing

Tiancheng Li ; Ninghui Li ; Jian Zhang ; Molloy, I. Knowledge and Data Engineering, IEEE Transactions on Volume: 24, Issue: 3 DOI: 10.1109/TKDE.2010.236 Publication Year: 2012, Page(s): 561 - 574

[5] Hsiao-Ying Lin; Tzeng, W.-G.; "A Secure Erasure Code-Based Cloud

Storage System with Secure Data Forwarding," Parallel and Distributed

Systems, IEEE Transactions on, vol.23, no.6, pp.995-1003, June 2012.

[6] C. Wang, Q. Wang, K. Ren, and W. Lou, “Ensuring Data Storage

Security in Cloud Computing,” Proc. 17th Int’l Workshop Quality of
Service (IWQoS ’09), pp. 1-9, July 2009.

[4] Amazon.com, “Amazon Web Services (AWS),” http://aws. amazon.com, 2009.

[8] Sun Microsystems, Inc., “Building Customer Trust in Cloud

Computing with Transparent Security,” https://www.sun.com/ offers/details/sun_transparency.xml, Nov. 2009.

[9] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, “Enabling Public

Verifiability and Data Dynamics for Storage Security in Cloud Computing,” Proc. 14th European Symp. Research in Computer
Security (ESORICS ’09), pp. 355-370, 2009.

[10] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, “Provable Data Possession at Untrusted Stores,” Proc. 14th ACM Conf. Computer and Comm. Security (CCS
’07), pp. 598-609, 2007.

[11] C. Aggarwal, "On k-Anonymity and the Curse of Dimensionality,"
Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 901-909, 2005.

[12] A. Blum, C. Dwork, F. McSherry, and K. Nissim, "Practical
Privacy: The SULQ Framework," Proc. ACM Symp. Principles of
Database Systems (PODS), pp. 128-138, 2005.

[13] J. Brickell and V. Shmatikov, "The Cost of Privacy: Destruction of
Data-Mining Utility in Anonymized Data Publishing," Proc. ACM
SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD),
pp. 70-78, 2008