IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Data Partitioning Technique to Improve Cloud Data Storage Security

Kiran Gabhale[1], Narendra Jadyal[2], Anurag More[3], Vinayak Bhalekar[4], Prof. V.V. Dakhode[5]

Department of Computer Engineering SKNCOE, SPPU, Pune, India[1,2,3,4,5]

Abstract: Cloud computing enables on-demand network access to a shared pool of conquerable computing resources such as servers, storage and applications. These shared resources can be rapidly provisioned to the consumers on the basis of paying only for whatever they use. Cloud storage refers to the delivery of storage resources to the consumers over the Internet. Private cloud storage is restricted to a particular organization and data security risks are less compared to the public cloud storage. Hence, private cloud storage is built by exploiting the commodity machines within the organization and the important data is stored in it. When the utilization of such private cloud storage increases, there will be an increase in the storage demand. It leads to the expansion of the cloud storage with additional storage nodes. During such expansion, storage nodes in the cloud storage need to be balanced in terms of load. In order to maintain the load across several storage nodes, the data need to be migrated across the storage nodes. This data migration consumes more network bandwidth. The key idea behind this paper is to develop a dynamic load balancing algorithm to balance the load across the storage nodes during the expansion of private cloud storage.

Keywords: TPA-Third Party Authentication, MD5-Message Digest 5, Data centre, Data Partitioning Technique.

I. INTRODUCTION

Cloud computing is an internet based technology, Cloud Computing is using hardware and software as computing resources to provide service through internet, Cloud computing being used widely nowadays to enable the end user to create and use software without worrying about the execution of the technical information from anywhere at any time. Over the network the resources are utilized and after computation these are delivered as services in cloud computing.

The Cloud Computing technology is embedded with three services which are just one click away, easy to use and pay as you use the service. Cloud storage is a service for developers to store and access data in cloud. Cloud service provider will manage and control the cloud resources. The benefits of the cloud storage are flexible with reduced cost and they also manage the data loss risk. many work focus towards third party auditing and the remote integrity checking, providing the data dynamics. Remote archive service is responsible for properly preserving the data.

The remote data integrity checking protocol detects the data corruption and misbehaving server in the cloud storage. In the proposed work Data partitioning technique, remote data integrity checking is analyzed in internal and external ways. Partitioning happens in alphabetical order by using of index method whereby the data being used is controlled. The security mechanism is also emphasized in order to prevent unrecoverable data loss. Storage and retrieval process are simplified by reducing the storage space when there is need to store and retrieved by merging technique.MD5 concept are used to check the integrity of data before storing of the data in dacenter.AES algorithm are used to store end user client data for security and RSA are used for communication of secure cloud data for storing and retrieving process.

II. RELATED WORK

In the Data Partitioning Technique literature review is done for data integrity checking and data storage mechanisms that are currently used in dynamic multi transactional applications. The dynamic data storage with token pre-computation and AES algorithm how it is stored in cloud is analysed; Integrity checking concepts is also used to detect and avoid misbehaving server considering data correction and error localization. Distributed scheme is used to achieve the data quality, availability, integrity of dependable storage services.

The data storage using dynamic data operation method is used to perform various operations. Security analysis is done by RSA to encode the data. Distributed storage system is also used to support the forwarded data in cloud without retrieval, ensuring secured and robust data in cloud storage.

Data integrity in cloud storage devices are analysed in the research works. Dynamic data operation 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. Storage model is also devised here to support multiple auditing tasks to improve efficiency.

In the works author considers generating signature methods for ensuring the cloud storage security.

Dynamic operations are up ported by using the RSA method. This method discusses data integrity and data correctness stored in cloud.

III. PROPOSED SYSTEM

In cloud data storage system, the clients stores data in cloud and also they maintain data locally.

Copyright to IJARCCE

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

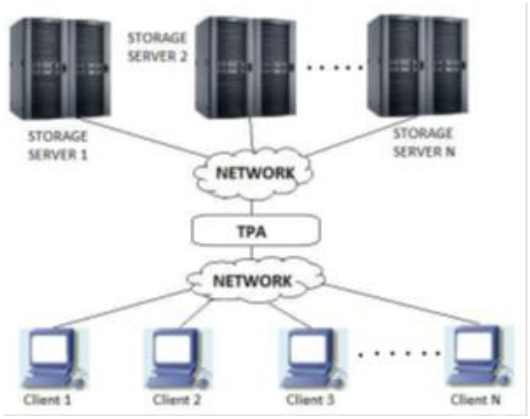


Figure 1. System Architecture

Here Partition of data provides security is in providing the security and avoid local copy of data. Fig.3 show the previous proposed system architecture. Proposed system divided in 3 different layers as follow:

1. **Client machine:** client machine are used by users who have to be data stored on cloud. Client machine either PC or browser enabled mobile device that rely on the cloud for data computation, consist of individual consumers and organizations.
2. **Cloud storage server:** Manage and provide storage space, computational resources and storage services by the cloud service provider (CSP).
3. **Third Party Auditor (TPA):** TPA, who has expertise and capabilities that users may not have, is trusted to assess and expose risk of cloud storage services on behalf of the users upon request, TPA perform security related operations. Partitioning data Partitioning of data performed at Third Party Auditor. Partitioning module accept user input file. Partitioning function has an important role in this work. It Splits (break up) larger files into smaller parts. It helps to store the data effectively in quick manner enhancing easy access to data also when there is need. The original data is complex and there is Difficulty in storing it in cloud, so partitioning function is used to make the storage easy in cloud. The partitioned files are encrypted, that is encoded with the key and stored in cloud. Partitioning takes place automatically when the data is fed for storing in cloud. Original file is also reconstructed when there is need to access the same

RSA Algorithm

RSA algorithm is designed by Designed by Ron Rivest, Adi Shamir, and Leonard Adleman Published in 1977.Most commonly used for encryption and authentication algorithm. It involves a public key and private key. The public key can be known to everyone and is used for encrypting messages. The basic steps of RSA algorithm are Key Generation Encryption and Decryption

Encryption:

Encryption technique is used to encrypt the files for security. By encrypting the file, the file will be in cipher. A common approach is used to encrypt with shared key algorithm and public key is randomly generated. Here we create public and private RSA key for encrypting the files, and stored in cloud. The generated private key length is 2048 bits. Secret key is symmetric encryption and public key is asymmetric encryption.

Algorithm 1: Encryption

1. Create a Cipher object and Key Generator object.
2. Create a Secret (private) key using cipher object.
3. Initialize it with private key.
4. Encrypt the files.
5. Get recipient's public key and Create Cipher and Initialize it for encryption with recipient's public key.
6. Create Sealed Object to seal session key using Asymmetric Cipher and Serialize Sealed Object.
7. Return the encrypted files and serialized Sealed Object to recipient.

Decryption:

Decryption technique is used to decrypt the files and the private key is generated to access files from cloud. For each end user separate private key is generated to access from any location with security. Non shared private key is used to decrypt files. The private key is an asymmetric technique. When decrypting files private key is generated for accessing ensuring file access control.

Algorithm 2: Decryption

1. Get encrypted message and serialized Sealed Object
2. Re-serialize Sealed Object. Create Cipher object, and initialize it for decryption
3. And generate private key.
4. Unseal the key using the asymmetric Cipher.
5. Create Cipher object and Initialize it with the Recovered private key for decryption.
6. Decrypt the files for access.

Partitioning Data

Partitioning function plays an important role in this work. 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 need. The original data is complex and there is difficulty in storing it in cloud, so partitioning function is used to make the storage easy in cloud. Partitioning happens alphabetical order by using index method.

It retrieves first two letters and checks it in folder with present having same letter. If it is not present then creates a folder and store the file in that folder .The partitioned files are encrypted, that is encoded with the Public key and stored in cloud. Partitioning takes place automatically when the data is fed for storing in cloud. Original file is also reconstructed when there is need to access the same. The partitioning concept is provided in the following algorithm.

Algorithm 1: Partitioning

1. Load the Input file with name.
2. Retrieve first two letters.
3. Check it in folder.
4. With present having same letter.
5. If not present then create a folder.

DOI 10.17148/IJARCCE.2016.5306

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

6. Encrypt all partition file with the help of public key and store the file in cloud data center
7. Decrypt the original file with the help of private key when there is need to access the end user.

MD5 Message Digest Algorithm

MD5 algorithm was developed by Professor Ronald L. Rivest in 1991. According to RFC 1321, "MD5 message digest algorithm takes as input a message of arbitrary length and produces as output a 128-bit "fingerprint" or "message digest" of the input.

The MD5 algorithm is intended for digital signature applications, where a large file must be "compressed" in a secure manner before being encrypted with a private key under a public-key cryptosystem such as RSA. MD5 is an algorithm that is used to verify data integrity through the creation of a 128 bit message digest from data input.

MD5 digests have been widely used in the software world to provide some assurance that a transferred file has arrived intact.

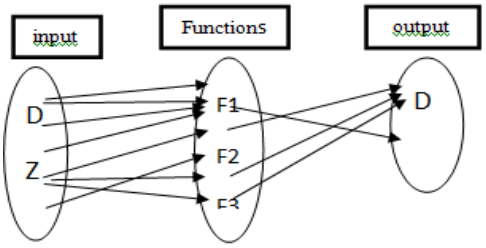
MATHEMATICAL MODEL

A. Let 'S' be the | Quick blood donate as the final set
S = {……
Identify the inputs as D, Z, N, and F}

B. Identify the inputs as D, Z, N, F
S = {D, Z, N, R, Q …
D = {D1, D2, D3…| 'D' gives Data to be stored or download from cloud}
Z = {Z1, Z2, Z3 …| 'Z' is the size of data}
N = {N1, N2, N3 …| 'N' is Number of clouds}
R = {R1, R2, R3 …| 'R' is fragmented data}
Q = {Q1, Q2, Q3 …| 'Q' is request to download data}

C. Identify the outputs as O
S = {D, Z, N, R, Q …
D = {D1, D2, D3…| 'D' gives Data to be stored or download from cloud}
R = {R1, R2, R3 …| 'R' is fragmented data}

D. Identify the functions as 'F'
S = {D, Z, N, R, Q, F…
F = {F1 (), F2 (), F3 (), F4 (), F5 (), F6 ()}
F1 (D):: Upload data
F2 (D, Z, and N): divide data into equal fragments compared to no of clouds
F3 (R): store data
F4 (Q): Request for download
F5 (R): Combine fragments of data
F6 (R): Download



IV. CONCLUSION

We propose an efficient data storage security in cloud service. The partitioning of data enables storing of the data in easy and effective manner. It also gives way for flexible access and there is less cost in data storage. The space and time is also effectively reduced during storage. Dynamic operation is another key 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 to provide higher level of security and searching mechanisms for outsourced computations in cloud services.

V. FUTURE WORK

Future work is planned to provide higher level of security and searching mechanisms for outsourced computations in cloud services.

REFERENCES

[1]. M. Armbrust et al., "Above the clouds: A Berkeley view of cloud computing," Dept. Elect. Eng. Computer. Sci., Univ. California, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2009-28, 2009.
[2]. G. Ateniese et al., "Provable data possession at untrusted stores," in Proc. 14th ACM Conf. Comput. Commun. Secur. (CCS), New York, NY, USA, 2007, pp. 598–609.
[3]. A. Juels and B. S. Kaliski, Jr., "PORs: Proofs of retrievability for large files," in Proc. 14th ACM Conf. Comput. Commun. Secur., 2007, pp. 584–597.
[4]. Anthony T.Velte, Toby J.Velte, Robert Elsenpeter, Cloud Computing A Practical Approach, TATA McGRAW-HILL Edition 2010.
[5]. Martin Randles, David Lamb, A. Taleb-Bendiab, A Comparative Study into Distributed Load Balancing Algorithms for Cloud Computing, 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops.
[6]. Mladen A. Vouk, Cloud Computing Issues, Research and Implementations, Proceedings of the ITI 2008 30th Int. Conf. on Information Technology Interfaces, 2008, June
[7]. Martin Randles, Enas Odat, David Lamb, Osama Abu- Rahmeh and A. Taleb-Bendiab, "A Comparative Experiment in Distributed Load Balancing", 2009 Second International Conference on Developments in eSystems Engineering.

Copyright to IJARCCE
DOI 10.17148/IJARCCE.2016.5306