# Secure Big Data Storage and Sharing Scheme for Cloud Tenants with Trapdoor Function

Mr. Vishnu Kumar Pandey<sup>1</sup> Dr. Kiran Bhandari<sup>2</sup>

<sup>1</sup>M. Tech. Student <sup>2</sup>Associate Professor

<sup>1,2</sup>Thakur College of Engineering and Technology, Mumbai-400 101, India

**Abstract**— Cloud computing is one of the most rapidly growing area which provides flexible, elastic and ondemand storage services for users. As cloud storage are growing rapidly data confidentiality, integrity and security in cloud storage system are always cause of concern. They are subject to attacks, modification and sometimes they even get stolen from storage system as our traditional security mechanisms doesn't provides enough security to our data. We are proposing new methodology where the mapping of big data are protected using trapdoor function and further we provide binary key encryption to every block of data.

**Keywords**— Cloud Computing, Big Data, Challenges, Issues, Solution

#### I. INTRODUCTION

Big data is set of data which are beyond the ability of commonly used software systems to store, manage, and process within a tolerable elapsed time, they are ranging from terabytes to many petabytes. A data centre mainly responsible for storing and processing of big data and those data are used for future scientific endeavors, Many IT industry are building their own data centre to accommodate huge data. Consequently, large amounts of data requires security while processing and modification, as these data are very crucial for any organization we need to provide security at micro level in cloud storage.



Fig. 1: Architecture of Typical Storage Environment

Cloud storage system consist of many licensing and delivery models such as Software as a service (SaaS), Platform as a service (PaaS) and Infrastructure as a service (IaaS), They provides different services to customers as service on demand. These data services as provided over internet to clients from cloud storage media at different places, Users can have access to data from different locations with security intact as they are getting accessed from different location. SaaS provides software services.

Software as a Service (SaaS) provides application to consumer which runs on cloud infrastructure and clients can access services from client side service like web browser or program interface, All services are controlled by service provider consumer do not have access to any cloud infrastructure.

Platform as a service (PaaS) is layer above IaaS which provides development tools and software hosted on provider's servers. This offers complete environment to developer for building their application, It provides complete software development life cycle from planning to design to building applications to deployment to testing to maintenance.

Infrastructure as a Service (IaaS) provides dedicated resources to clients, this reduces the cost for huge initial investment.

proposed scheme, In the Data blocks (D1,D2,...,DN) are divided into multiple blocks (B1,B2,...B3) of messages where each block have different data which will be stored in cloud storage. When data blocks are stored in cloud they will get encrypted using encryption keys (K1,K2,...,K3), After storing the data blocks can be shared with authorized users.



Fig. 2: Data in cloud storage system with Trapdoor Function

#### II. LITERATURE SURVEY

Data security is practice to protect data from unauthorized access and ensure privacy while protecting personal or corporate data. In a distributed environment, datasets are located in different data centre and therefore face challenges such as data security, privacy protection and authentication, Many scheme have been proposed in past but they cannot be applied on big data efficiently as those algorithms are costly in terms of time and space.

The first scheme is encryption which uses mathematical algorithms to scramble data into unreadable text. It can only be decrypted by the user who have valid decryption key, this method uses full file-level encryption, this scheme provides good security but it introduces some challenges with respect to key management and thus causes low efficiency and more time consumption.

The second scheme is application-level data encryption technology, This technology provides security when there is vulnerability in network level, it ensure that only certain users get to access the data through a particular

7

application. This scheme will be very costly because it must maintain many parameters and data structures.





The final scheme is user authentication which can be applied to the application, files, folder and on computer system. This method provides various privileges when you logged in, Some system will force you to re-enter password if machine is been idle for some time and you have to provide authentication once again. However, this method is quite popular but in big data system it leads to decline in system performance as there are huge amount of data available to get encrypted, As security is entirely based on confidentiality and the strength of the password and it may possible that there is lack of security and identity check.

# III. PROPOSED FRAMEWORK

As traditional security schemes are not efficient to protect big data we have come up with solution that prevents unauthorized access of data and meantime it will give secure access to the users as well. Today most of the user data are stored in cloud platform, People have variety of concern when they store their confidential on cloud storages. Undoubtedly, Privacy and security of personal data information is the most important concern for users.

In order to make the big data of our users secure, we propose a secure cloud based storage system with binary encryption with trapdoor function. Each dataset will be separated into a sequence of n parts, where each part can be denoted by part i(i (1, n)), and they will be stored at m different storage providers, where each provider is identified as provider j( j (1,m)). These m storage providers may belong to one or more storage providers. so, when big data of a users are stored, it will form unique storage path for the big data given as MappingStorage\_Path={Data.((P1(M1,M2 ... Mr) (P2(M1,M2 ... MS); ... (Pn(M1,M2 ... Mt)); where, P denotes the storage provider, and M denotes the physical storage media.

As big data are enormous and impossible to encrypt them as whole we will encrypt the storage path of big data with key an called virtual mapping of big data. This is the system which protects the mapping of the various data elements to each provider using a trapdoor function, the proposed scheme will distribute all data in different storage service providers. In case of data lost the owner have the index information of each data parts, So, when some data parts is on cloud storage provider lost, we can try to find another copies of the data parts according to their storage index information.

Our scheme have modal in which data will be stored in cloud storage along with keys with key manager. The first layer is storage model where data will be stored in server through cloud storage provider, Second layer have functionalities of key management where we initialize the keys of various data blocks of system. Third layer in which the server will get encrypted with symmetric encryption keys, from server the data can be accessed by different users where these data will be shared among different users.



Fig. 4: Encryption key model in Cloud Storage

In the proposed scheme, our big data will be separated into multiple sequenced parts and then will be stored on different media. When users want to access their data then data from different part will be collected and clients gets access have data as one. these is very crucial because this protect our data from getting stole by any means, Further we use more advance security mechanism to protect our data. These data are classified into public data and confidential data.

We are introducing trapdoor function with binary encryption but prior to describing our proposed scheme we will introduce trapdoor function. A trapdoor function is a function that is easy to compute in one direction but difficult to compute in other and some piece of information can be made hard direction much easier. This can be explained as A can easily compute the encryption of his message using B's public key, but it will be very hard for C to reverse this process. B can use private key to read A's message.

Trapdoor function will provide strong security to our system along with that we propose proxy re-encryption scheme, a proxy re-encryption schemes allows third parties to alter cipher text which is encrypted by one party, so that it may be decrypted by another. These scheme are applied when user want to share his data at that time he have to send re-encryption key to the storage server after that server will re-encrypt the message and for authorized users this will increase data confidentiality and enhances the data forwarding function.

Cloud computing emerged as best data management service for efficient storage and economic data processing model. Many companies or enterprises face several difficulty as they have to do maintenance storage and management by themselves or find reliable institutions, meanwhile transmission and processing time is too long and there are chances of system crash or failure remains high, in the proposed scheme, The big data of users will be divided into smaller data blocks, these smaller data blocks will be stored in cloud media storage which increase the efficiency in data transmission and storage. Our proposed scheme have lower transmission failure probability and mainly focused

upon sharing of confidential data, our scheme protect the privacy and confidentiality of data and they are remains non-accessible to unauthorized users.



Fig. 5: Delivery Mechanism in Cloud storage model

Cloud storage provides variety of services when users store data on single storage device it may have several drawbacks such as service error, System crash then users have chances of losing data or data unavailability. In our proposed scheme, data are stored in the form of block form and as we are using redundant data backup strategy, In case if storage service failed, then user will not be affected therefore we avoids sudden loss of data. In the proposed scheme, Storage path of big data is encrypted to prevents confidentiality and unauthorized access and we can call cryptographic value as virtual mapping of big data this is very efficient system to protect data as it is in almost k level and sharing with other users by distributing the secret information with identity encryption.

# IV. SIMULATION

To describe the simulation we design a procedure to establish communication key between different users, which used to establish communication among user for sharing bid data. We use different parameters to describe the identity element of system, Zq to denote group under modulo q.

# A. Algorithm

The procedure of establishing communication key between user A and other cloud user, for example B Tenant A computes parameter  $Y_A$ : Choose  $X_A < q$  $Y_A = \eta^{XA} \mod q$ . User B computes parameter  $Y_B$ : Choose  $X_B < q$ Compute  $Y_B = \eta^{XB} \mod q$ . Tenant A encrypts  $Y_A$ ,  $ID_A$  and  $ID_B$  using IBE algorithm and then send to B: Encrypt (YA, IDA and  $ID_B$ ) $\rightarrow$ B User B encrypts  $Y_B$ ,  $ID_B$  and  $ID_A$  using IBE algorithm and then send to B: Encrypt (YA, IDA and  $ID_B$ ) $\rightarrow$ A Tenant A decrypts message and computes  $K_1 = (Y_B) X_A \mbox{ mod } q$ 

User B decrypts message and computes  $K_2 = (Y_A)X_B \mod q$ From the above definitions, we can conclude that  $K_1 = (Y_B)^{XA} \mod q = \eta X_A X_B \mod q K_1 = (Y_B)X_A \mod q = X_A X_B \mod q = (Y_A)X_B \mod q = K_2$ .





In Simulation, we evaluate the efficiency and data storage robustness of proposed scheme and traditional scheme using different scenarios. In first simulation scheme, big data transmission overhead of traditional scheme will be compared with proposed scheme with same data size; we test with these condition 5 times and the results are shown in diagram. In second simulation we check the data availability in serving data to cloud clients, the traditional schemes may have unavailable during failure time but newly proposed scheme will be robust as data are divided into multiple parts. we have stored data on cloud servers and tested these condition for five times with different probability and the results are displayed in the graph.



Fig. 7: Data availability of two models

# V. CONCLUSION AND FUTURE WORK

In this paper, we proposed new method for constructing Secure Big Data Storage and Sharing Scheme for Cloud Storage using trapdoor function. as big data requires large space and time our proposed model avoids this by splitting data into multiple blocks and protect it with trapdoor function. We analyzed the proposed scheme with respect to security and efficiency and all results show that proposed scheme is effective and feasible to protect the big data. Future researches might consider in future such as use of biometric signatures to authenticate and encrypt data.

#### REFERENCES

 Hsiao-Ying Lin, Member, IEEE, and Wen-Guey Tzeng, "A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding", IEEE Transactions On Parallel And Distributed Systems, Vol. 23, No. 6, June 2012.. [2] X. Zhang, H. tao Du, J. quan Chen, Y. Lin, and Ljie Zeng, "Ensure data security in cloud storage," in Network Computing and Information Secu - rity (NCIS), 2011 International Conference on, vol. 1, may 2011, pp. 284 –287.Data Integrity, from Wikipedia, the free

encyclopedia,[Online]https://en.wikipedia.org/wiki/Big \_data.

[3] Liu Q, Wang G, Wu J. Secure and privacy preserving keyword searching for cloud storage services [J]. Journal of network and computer applications, 2012, 35(3): 927-933.