IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 9, Issue 5, May 2020

# Intelligent Data Backup Technique for Cloud Storage

**Ms. Namita Jakulwar[1], Mr. Hirendra Hazare[2]**

M. Tech Student, Department of CSE, Ballarpur Institute of Technology (BIT), Ballarpur[1]

Assistant Professor, Department of CSE, Ballarpur Institute of Technology (BIT), Ballarpur[2]

**Abstract:** In cloud computing, data generated in electronic form are in large amount. To maintain this data efficiently, there is a necessity of data recovery services. To cater this, in this topic we propose a smart remote data backup plan using Seed Block Algorithm (SBA) with Advance Encryption Standard (AES) algorithm. In this topic we are proposing a procedure which allows users to store their data onto the cloud, as soon as the file is stored at the first cloud server it gets encrypted using AES. In case if the certain file gets deleted due to any reason, AES helps to recover that file from a backup file which is stored at a remote location. The time related issues are also being solved by proposed method such that it will take minimum time for the recovery process. Proposed method also focuses on the security concept for the back-up files stored at remote server using AES encryption algorithm.

**Keywords:** Seed Block Algorithm, AES, Cloud back-up, Remote cloud, Main Cloud.

## I. INTRODUCTION

Cloud Computing is itself a gigantic technology because of its advantages over previous systems like grid or cluster computing. Cloud storage provides online storage where data stored in form of virtualized pool that is usually hosted by third parties. Number of users share the same cloud storage provided by a certain service provider.

Faulty equipment's, a human error, network connectivity, a bug or any criminal intent may put our cloud storage's security at stake. Cloud service provider may also make some changes in the configuration; this may lead to loss of alteration of the information stored by user. There is possibility of data loss. To solve these difficulties, we need to provide data integrity for our cloud. In literature many techniques have been proposed PCS [1], HSDRT [2], Linux Box [3], ERGOT [4], Cold/Hot backup strategy [5] etc. that, discussed the data recovery process. However, still various successful techniques are lagging behind in some critical issues like implementation complexity, low cost, security and time related issues.

To overcome the disadvantages of previously proposed systems we have proposed and are implementing a new method based on Seed Block Algorithm (SBA) and Advance Encryption Standards (AES) Algorithm. The mentioned procedure works in following manner: in first step it allows users to collect and store their files onto the main cloud. As soon as the files get stored at the cloud, those get encrypted using AES algorithm. In step two, in case of file deletion it helps user to recover the files.

## II. LITERATURE REVIEW

In literature survey, we have studied the most recent back-up and recovery techniques that have been developed in cloud computing domain such as PCS [1], HSDRT [2], Linux Box [3], ERGOT [4], Cold/Hot backup strategy [5] etc. When we studied the existing methods in detail, we found that, performance of the system is not satisfactory with respect to cost, security, low implementation complexity, redundancy and recovery in short span of time.
We inferred after study of various present techniques that PCS is comparatively reliable, simple, easy to use and more convenient for data recovery totally based on parity recovery service. It has higher probability and efficiency of recovering among present techniques. It generates a virtual disk in user system for data backup, make parity groups across virtual disk, and store parity data of parity group in cloud to recover the data. It makes use of the Exclusive OR functionality for creating Parity information. However, there are some problems associated with this method. This method is unable to control the implementation complexities.
On the other side, HSDRT method ensures as a powerful technique for the movable clients such as laptop, smart devices, palmtops etc. However, it is not economical for the implementation of the recovery and also unable to control the data replication. It an innovative _le back-up concept, which makes use of an effective ultra- widely distributed data transfer mechanism and a high-speed encryption technology.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 9, Issue 5, May 2020

The HS-DRT [2] is an innovative file back-up concept, which makes use of an effective ultra-widely distributed data transfer mechanism and a high-speed encryption technology. This system follows two sequences one is Backup sequence and second is Recovery sequence. In Backup sequence, it accepts the data to be backed-up and in Recovery Sequence, when some calamities occur or periodically, the Supervisory Server starts the recovery sequence. However, there are some limitations in this model and therefore, this model somehow fails to declare as perfect solution for back-up and recovery.

## III. PROPOSED SYSTEM

The Seed Block Algorithm is used in the proposed system so as to ensure a secure backup of data at the main cloud and on the remote server as well. Even if main cloud gets crashed / damaged or by mistake the files on it get deleted, as the backup of such files are stored at remote server, the owner of respective file obtains the original files again from remote server. Proposed SBA also focuses on the security concept for the back-up files stored at remote server, it has low implementation complexity and the time related issues are also being solved by proposed SBA such that it will take minimum time for the recovery process. AES, which has overcome the disadvantages of the RSA algorithm, is used for file encryption-decryption purpose. It is used to provide authentication to the file to maintain their confidentiality and integrity. Architecture of proposed system is shown in figure1

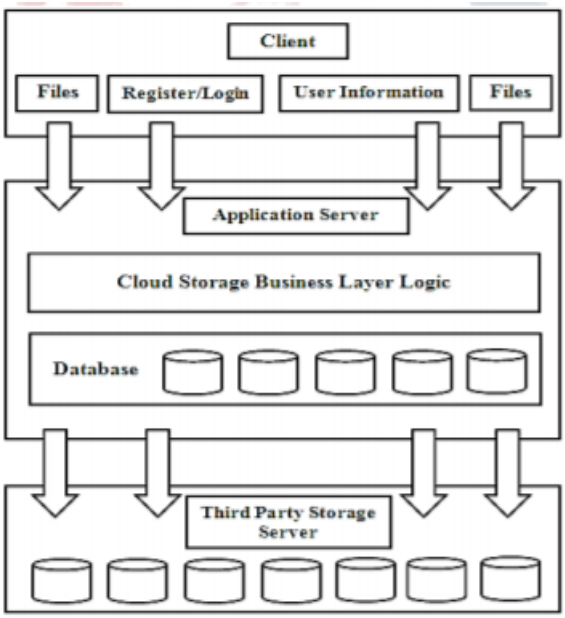



Fig 1: Architecture of proposed data backup cloud system using SBA security with AES

### A. Remote Data Backup Server

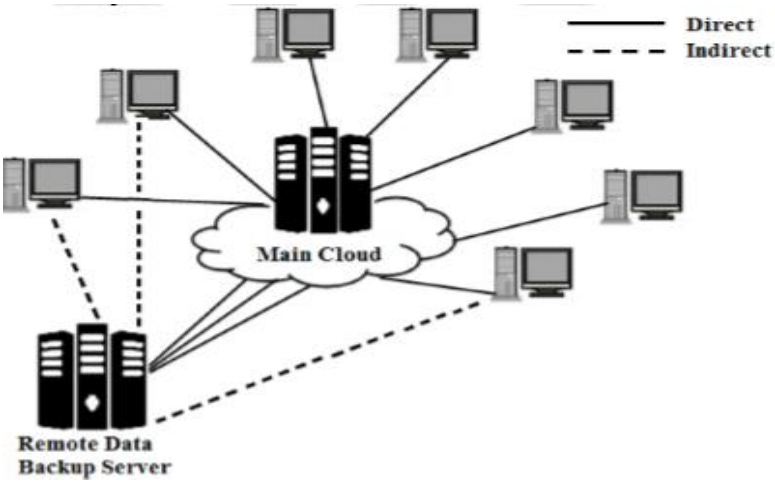



Fig 2: Remote Data Backup Server

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 9, Issue 5, May 2020

Backup server of main cloud keeps the copy of main cloud. However, when it is at remote location and possess the entire state of the main cloud, it's called as Remote Data Backup Server. Main cloud is called as central repository while remote backup cloud is called as remote repository. However, due any reason like any natural disaster (E.g. earthquake, flood, fire, etc.) or by human attack or file deletion done mistakenly, if the central repository lost its data then it utilizes the respective services for getting information from the remote repository. Information gathering for user from any remote location irrespective of network connectivity, even data cannot be retrieved from main cloud; these are the key goal of the remote backup. Fig.2 indicates that if data is not found on central repository then the access to the files is granted to users to from remote repository (i.e. indirectly) [8]. Remote backup services should cover issues like Data Integrity, Data security, Data Confidentiality, Trustworthiness, Cost efficiency.

The proposed system supports files of all extensions like .bmp, .gif, .png, .jpg, .jpeg for images, .txt, .doc, .docx, .xls, .xlxs, .pptx for textual data, .pdf of any combination of images and textual data. The user can even upload zip , rar files also audio files of various extensions like .mp3, .wma, .wav and video files of various extensions like .3gp, .avi, .flv, .mkv, .mov, .mp4, .wmv. A set of files can be uploaded at once by compressing the files using .zip extension. Also file sharing among all the users of drive is implemented which will help different users to share files with other users on the drive. Original file is given as input to SBA generating EXORed (□) file which is further given to the file encryption-decryption process is shown in figure 3. AES algorithm is used for encrypting & decrypting the files to provide the security. File sharing among all the users of drive is implemented. All the files that are shared will have Read-Only permission to all the users. Also, the author of file will have Read, Write and Execute permissions, by default and he can assign privilege permissions to a specific no. of users as desired.

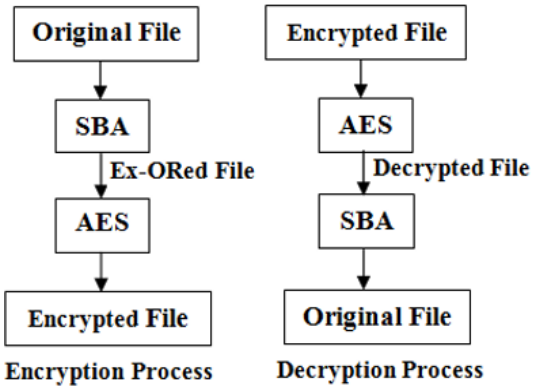


Fig 3: File Encryption-Decryption Process

**B. Seed Block Algorithm (SBA)**

Providing ease in the recovery process and back-up are the main objectives of Seed block algorithm. Exclusive-OR (XOR) operations are used in SBA. First for every client unique client id and a random number the cloud is set. Then, whenever any client registers in the main cloud then the random number and registered client_id get EXORed to generate seed block id for that specific client. The generated seed block id corresponds to each client is stored at remote server. In the cloud, At first time, client creates the file which is then stored at the main cloud server. Such main file of client is being EXORed with the Seed Block of the particular client. And that EXORed file is stored at the remote server in the form of file‟ (pronounced as File dash). If main cloud gets crashed / damaged or by mistake the files on it get deleted, as the backup of such files are stored at remote server, the owner of respective file obtains the original files again from remote server by EXORing file‟ with the seed block of the corresponding client. Proposed SBA algorithm is as follows:

**Initialization:**
Main Cloud: $\mathbf{M_c}$;
Remote server: $\mathbf{R_s}$;
Clients of Main Cloud: $\mathbf{C_i}$;
Files: $\mathbf{a_1}$ and $\mathbf{a_1'}$;
Seed block: $\mathbf{S_i}$;
Random Number: **r**;
Client's ID: $\mathbf{Client\_Id_i}$;
**Input:**
$\mathbf{a_1}$ created by $\mathbf{C_i}$; **r** is generated at $\mathbf{M_c}$
**Output**:
Recovered file $\mathbf{a_1}$ after deletion at $\mathbf{M_c}$

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 9, Issue 5, May 2020

**Given**:
Authenticated clients could allow uploading, downloading and do modification on its own the files only.
**Step 1:** Generate a random number
**int r = rand( ).**
**Step 2:** Create a seed Block $S_i$ for each $C_i$ and Store $S_i$ at $R_s$,
$S_i = r \oplus ient\_Id_i$
(Repeat step 2 for all clients)
**Step 3:** If $C_i$ /Admin creates/modifies $a_1$ and stores at $M_c$, then
$a_1$' is created as:
$a_1' = a_1 \oplus S_i$
**Step 4:** Store a1" at Rs.
**Step 5:** If server crashes a1 is deleted from Mc, and we do EX-OR to retrieve the original a1 as:
$a_1 = a_1' \oplus S_i$
**Step 6:** Return a1 to Ci.
**Step 7:** END. [9]

### C. Advanced Encryption Standard (AES)

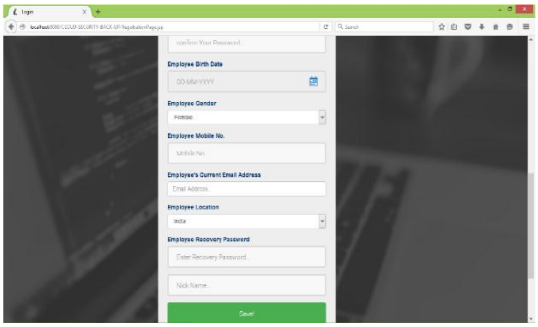
The Advanced Encryption Standard (AES) is a symmetric key encryption standard. The standard consists of three block ciphers AES-128 AES-192 and AES-256 adopted from Rijndael. Each of these ciphers has a 128-bit block size with key sizes of 128, 192 and 256 bits respectively. The key size used for an AES cipher denotes the number of repetitions of transformation rounds that converts the input called the plaintext into the final output called as cipher text. After generating the cipher text from the plain text, the encrypted file is stored on the main cloud and that file is only getting backed-up on the remote server.

## IV. EXPERIMENTAL RESULTS

Few snapshots of implementation modules:

a) Registration Page:
The user has to first register self for logging in into the drive. After which the user will get a proper username and a password. A seed_id is created for that individual user which will be used as a unique id for recognizing the user. The details will include all his personal information including his email address on which all the keys required to view or recover files will be sent.


b)



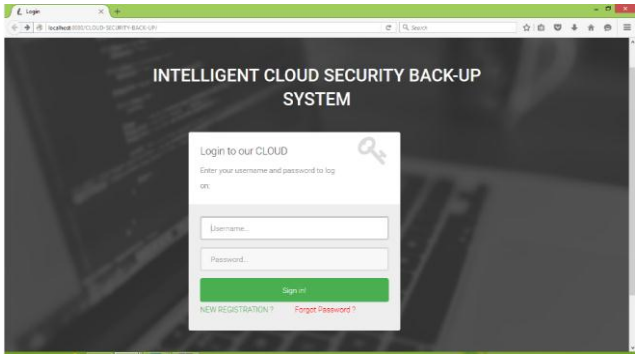
IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 9, Issue 5, May 2020

c) Login Page:




## V. RESULT & FUTURE SCOPE

This project is mainly indented to take Backup of data in LAN and WAN. As future expansion it can be implemented in World Wide Web. Another feature that can be implemented is incremental backup, which will enable to save storage space by smart backup. New technologies like Cloud Computing can be used to make the backup system more effective and efficient.

## VI. CONCLUSION

Disasters both natural and human-caused can threaten your precious files at any time: a fire, power surge, or leaking pipe could fry your system. Even without suffering a calamity, there are plenty of other threats to locally stored data hard drive failure, accidental erasures, or a lost or stolen laptop could make you a victim of data loss. By data, here, we mean things like your irreplaceable family photos, videos, and music as well as documents. Secure Network Backup System, securely store your files away from your premises at onsite -site server locations, your data will stay intact and available even if your local disks are stolen or your premises suffer some disaster. With more and more emphasis on "cloud computing," it only makes sense that backup should take advantage of this hot trend in technology.

## REFERENCES

[1] Kolipaka Kiran, Janapati Venkata Krishna, 2014, “Smart Data Back-up Technique for Cloud Computing using Secure Erasure Coding”, IJCTT–volume 16 number 3 – Oct 2014.
[2]. Ms. Kruti Sharma, Prof K. R. Singh, 2012, “Online data Backup and Disaster Recovery techniques in cloud computing: A review”, JEIT, Vol.2, Issue 5.
[3]. Tanay Kulkarni, Krupali Dhaygude, Sumit Memane, Onkar Nene, 2014, “Intelligent Cloud Back-Up System”, International Journal of Emerging Engineering Research and Technology, Volume 2, Issue 7, October 2014, PP 82-89.
[4]. Giuseppe Pirr´o, Paolo Trunfio, Domenico Talia, Paolo Missier and Carole Goble, 2010, “ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures”, 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.
[5] Vijaykumar Javaraiah Brocade Advanced Networks and Telecommunication Systems (ANTS), 2011, Backup for Cloud and Disaster Recovery for Consumers and SMBs, IEEE 5th, International Conference
[6] Giuseppe Pirro, Paolo Truno , Domenico Talia, Paolo Missier and Carole Goble, ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures, 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.
[7] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, Recovery Strategies for Service Composition in Dynamic Network, International Conference on Cloud and Service Computing.
[8] http://www.eecs.berkeley.edu/Pubs/ TechRpts /2009// EEC S-2009-28.pdf.