ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

Efficient Data Backup Technique for Cloud Storage

[1] Yogesh Gite, [2] Ankush Pawar, [3] Dr. Shashikant Ghumbre

Abstract: **- Large amount of data in electronic form are generated in cloud computing. Data recovery mechanisms are required to preserve such data proficiently. To cater such issue, in this paper we propose an efficient data backup technique Seed Block Algorithm (SBA) for cloud storage and securing back-up files stored at remote server with Advance Encryption Standard (AES) algorithm. In this paper a technique is proposed that permit users to store their data especially onto the first cloud server, once the file is stored at cloud server the AES technique encrypts such file. Due to any reason if any file gets deleted, the SBA along with the help of AES get back that file from remote location where the backup files are stored.**

Keywords:- **Seed Block Algorithm, AES, Encryption, Decryption, Cloud computing, Remote Data Backup Server.**

I. INTRODUCTION

To provide convenient, on-demand network access NIST, National Institute of Standard and Technology, describe a model that assist to a share various types of configurable computing service (for applications, services, storage, servers, ex-networks) that can be demanded rapidly and quickly released with less effort by services provider or the management. In the tough & competitive world of information Technology, Cloud Computing is overwhelming the recent technologies namely grid, cluster, distributed etc. The advantages of cloud computing conquered many disadvantages of many previous computing techniques to make it as gradually rising as an essential technology. The hosting company operates large data on large data center which are virtualized the resources as per the requirements of the customer and expose them as the storage pools that assist user to store files or data objects. [1] Various user shares their storage and other resources, also other users may require accessing the data you shared. The cloud storage gets risky may get in danger due to network connectivity, faulty equipments, a bug, human error or any criminal intention. On cloud storage the frequently changes are made, is called as data dynamics that supports many actions like insertion, deletion and block modification. As the services are not only restricted for operations like archiving and getting data backup but also the remote data integrity is also desired. The huge amount of data is generated on cloud that should remains unchanged during the process of transmission and storing data at main cloud remote server, as the data integrity of server usually concentrating on the validity and fidelity of the server's entire state. In the back-up and recovery services, integrity plays a vital role and file encryption process is carried out to maintain the integrity of data the stored on the Drive.

II. EXISTING SYSTEM

Cloud computing is required to make elegant data backup, since the need of cloud computing is rising continuously as advantages of cloud computing overwhelm disadvantages of recent computing technique like grid, cluster, distributed [1].

Sr. no.	Approach	Advantage	Disadvantage
1	Parity Cloud Service(PCS)	▪ Reliable ▪ Privacy ▪ Low cost	▪ Implementation complexity is high
2	HSDRT	▪ Useful for movable clients such as laptops and smart phones.	▪ Costly ▪ Increased redundancy
3	ERGOT	▪ Privacy	▪ Time complexity ▪ Implementation complexity
4	Linux Box	▪ Simple ▪ Low cost of implementation	▪ Requires high bandwidth ▪ Privacy ▪ Complete server backup at a time
5	Rent out rented resources	▪ Virtualization, rents it to the clients in the form of a cloud ▪ Cost depends on infrastructure utilization	▪ Implementation gets complex ▪ Resources must be kept under special attention due to rented concept
6	Cold/ Hot backup strategy	▪ Triggered only when failure is detected	▪ Cost increases as data increases gradually

Table I: Comparison between various techniques of backup and recovery [13]

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

In literature, most of recent backup and recovery techniques such as HSDRT, PCS, ERGOT, Linux Box and Cold/ Hot Backup Strategy are studied. Table-I shows detail review, under all uncontrolled circumstances such as redundancy cost, recovery, security, and low implementation complexity none of above techniques able to give best performance in short span of time. Even the existing implementations of recovery technique using Seed Block supported only the .txt and .docx formats. Also the user was not able to upload a set of files at once. In all previous implementations of Seed Block algorithm only the RSA algorithm was used for encryption-decryption purpose.

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 get 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 obtain 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

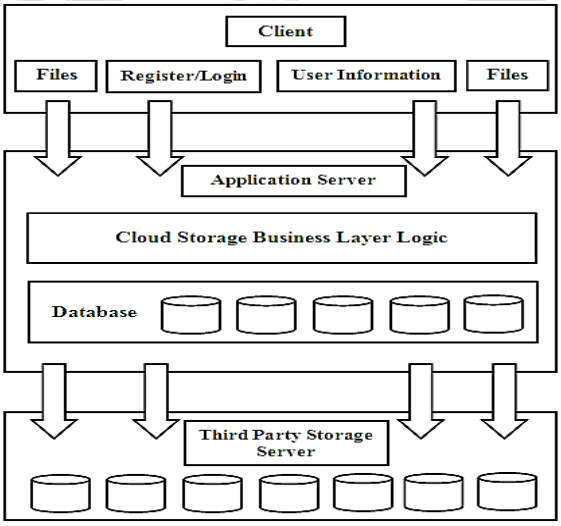


Fig 1: Architecture of proposed data backup cloud system using SBA security with AES

A. Remote Data Backup Server

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.

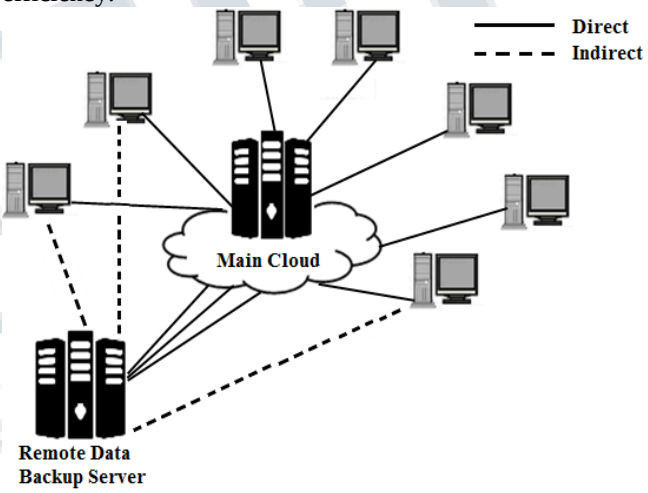


Fig 2: Remote Data Backup Server

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

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

file will have Read, Write and Execute permissions, by default and he can assign privilege permissions to a specific no. of users as desired.

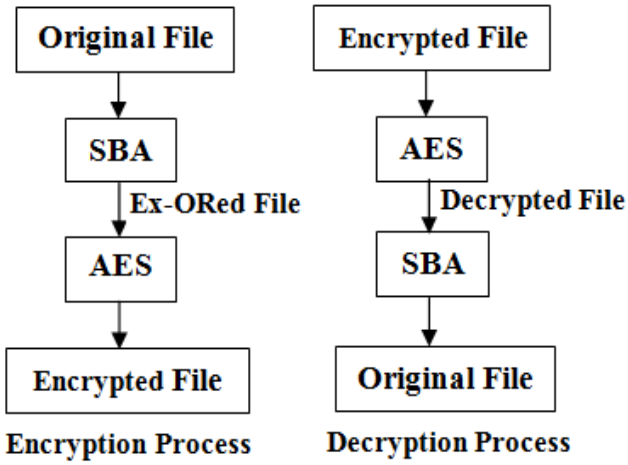


Fig 3: File Encryption-Decryption Process

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 are 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 get 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 obtain 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: **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 $\mathbf{M_c}$

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 $\mathbf{C_i}$ and Store $\mathbf{S_i}$ at $\mathbf{R_s}$,

$\mathbf{S_i = r \oplus Client\_Id_i}$

(Repeat step 2 for all clients)

Step 3: If C_i /Admin creates/modifies $\mathbf{a_1}$ and stores at $\mathbf{M_c}$, then

$\mathbf{a_1'}$ is created as:

$\mathbf{a_1' = a_1 \oplus S_i}$

Step 4: **Store** a_1' **at** R_s.

Step 5: If server crashes a_1 is deleted from $\mathbf{M_c}$, and we do EX-OR to retrieve the original $\mathbf{a_1}$ as:

$\mathbf{a_1 = a_1' \oplus S_i}$

Step 6: Return $\mathbf{a_1}$ to $\mathbf{C_i}$.

Step 7: END. [9]

C. Advanced Encryption Standard (AES) Algorithm

Flowchart: The following figure 4, flowchart describes the flow of the proposed method which contains the encryption and decryption processes [14].

Architecture: The proposed method is shown in the figure 5 in which the confidential data collected from the user will be divided based on size of the data. The data which is of large size can use 128bits key encryption. So that the data encryption time will be short as 128bits key uses 10 no of iterations. The data which is of medium size can use 192bits key encryption which is using 12 no of iterations. The data which is of small size can use 256bits key encryption which is using 14 no of iterations [14].

The encrypted data will be stored in the cloud storage. The user will be provided an authentication to access the data. He can download the data from the cloud. The data is encrypted with 128, 192 or 256 bits of key size based on the size of the data to be encrypted. The data can of any format like pdf, text files, images, videos, audio files, etc. The working of AES technique is shown in following figure 6.

AES the recent standard for secret key encryption which uses an iterative approach instead of using Feistel cipher.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

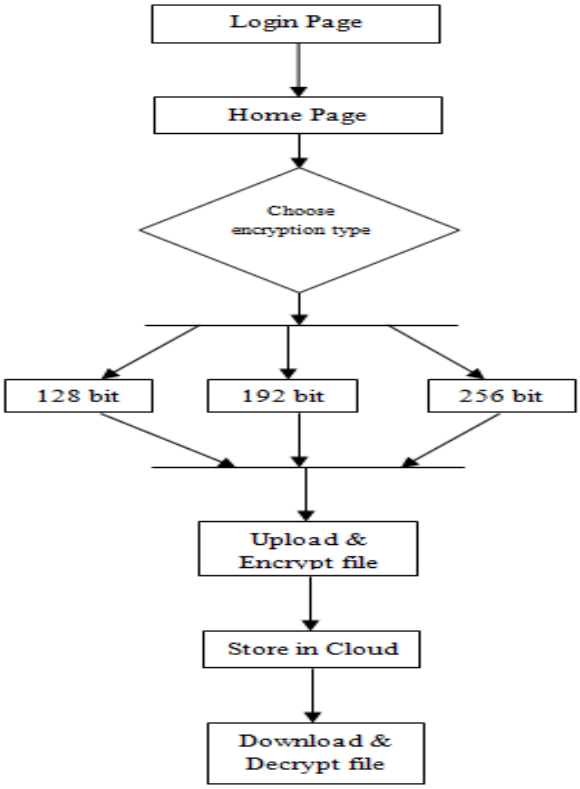


Fig 4: Flowchart of the proposed method

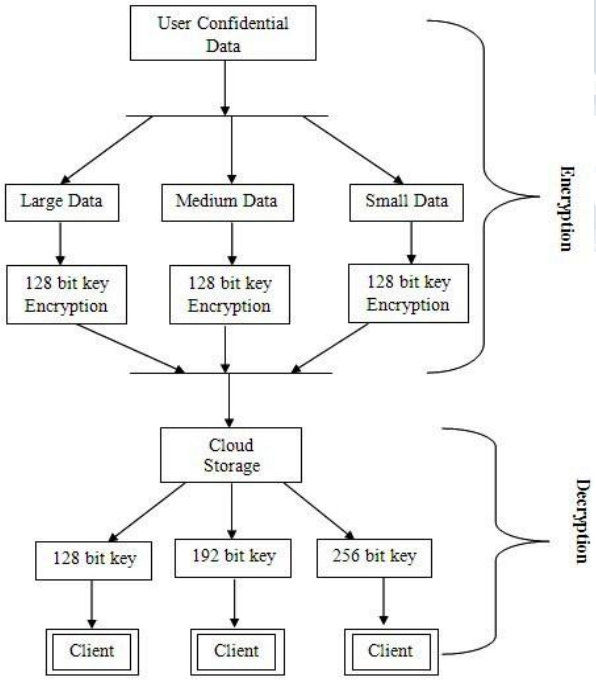


Fig 5: Architecture of AES Technique Implementation

It is based on 'substitution–permutation network'. The algorithm uses a combination of Exclusive-OR operations (XOR), octet substitution with an S-box, row and column rotations, and a Mix Column. It is easy to implement and run in an adequate amount of time on a normal computer. AES performs all its computations on bytes rather than bits so it considers 128 bits of a plaintext block as 16 bytes which are arranged in four columns and four rows for processing as a matrix. The number of rounds depends on the length of the key, it uses 10, 12, 14 rounds for 128-bit, 192-bit and 256-bit keys respectively. Each of these rounds uses a different 128-bit round key, which is calculated from the original AES key

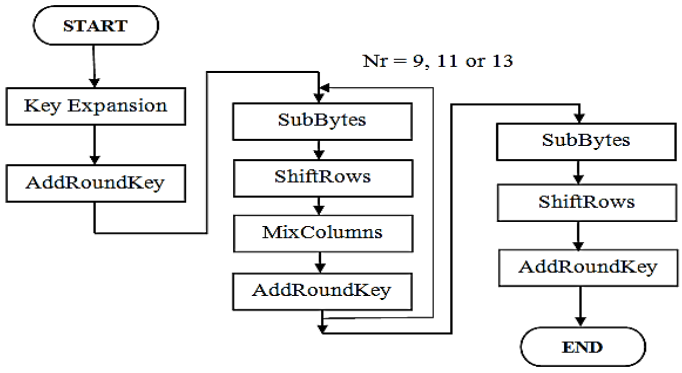
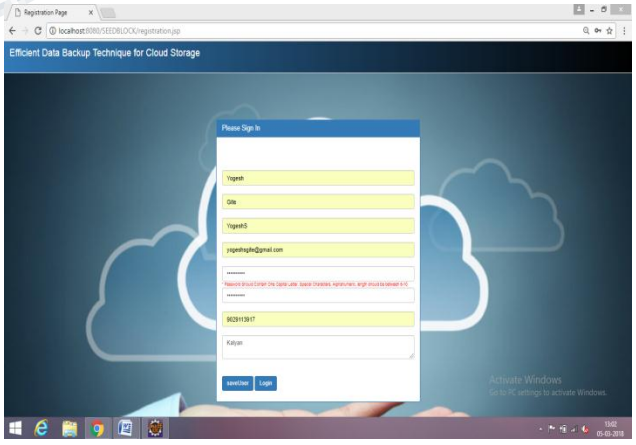


Fig 6: AES algorithm working

D. Modules

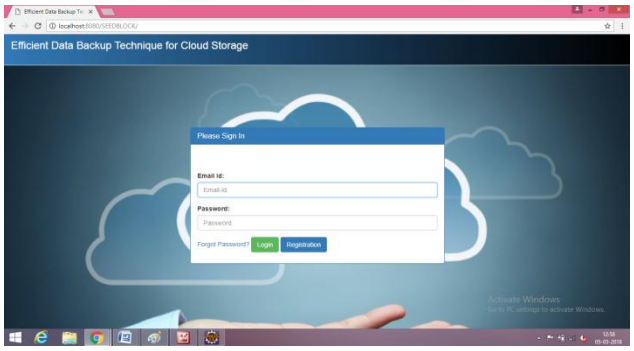
1. Registration and User login:

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.

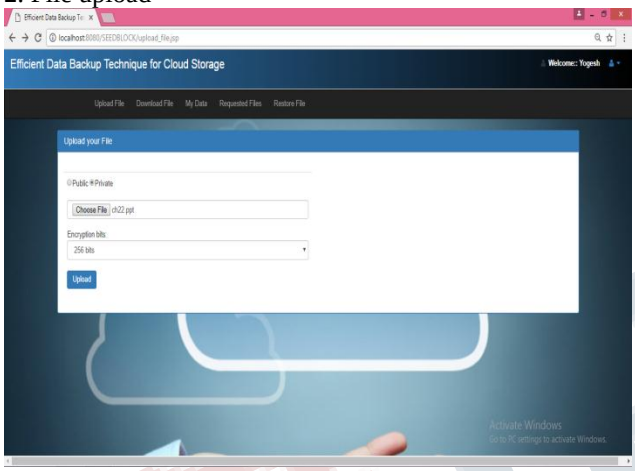


ISSN (Online) 2394-2320

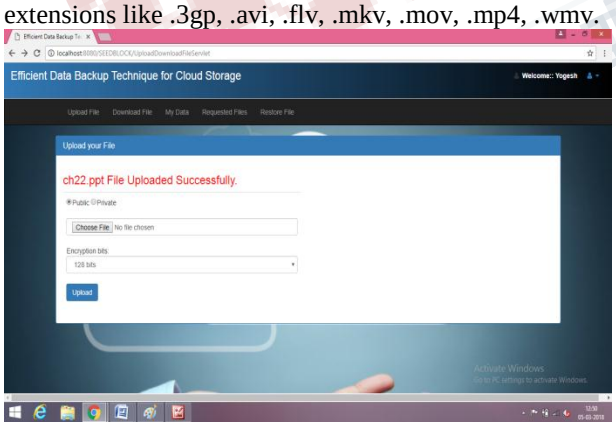
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018



2. File upload

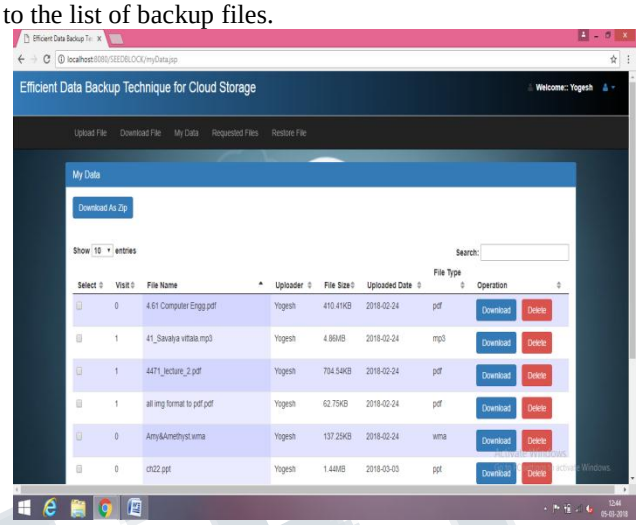


The user can upload files into his account through this module. User can upload files of various extensions like like .bmp, .gif, .png, .jpg, .jpeg for images, .txt, .doc, .docx, .xls, .xlxs, .pptx for textual data, .pdf of any combination of images and textual data, .zip, .rar files, audio files having extensions like .mp3, .wma, .wav and video files having extensions like .3gp, .avi, .flv, .mkv, .mov, .mp4, .wmv.

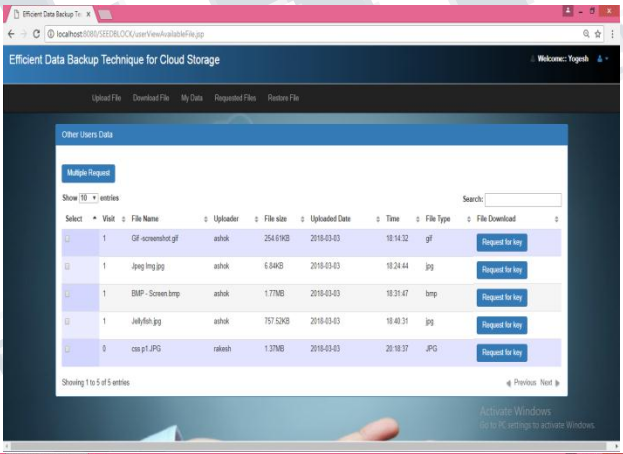


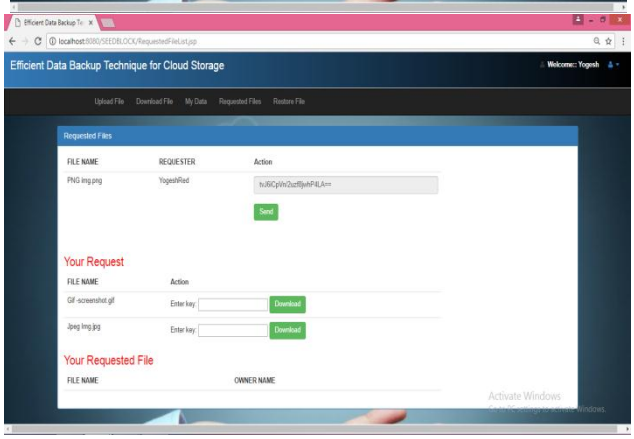
3. View files:

In this phase, the user can view all the files he has uploaded so far, along with the file's id, name and its type. If user wants to view the files he has uploaded, then he needs to first download it. In order to avoid the shoulder surfing attack, provide a One Time Password authentication system. OTP has been use in order to provide easy and fast security, which is sent to the user's email id. The user can also delete the file if it is no more needed and the deleted file is added to the list of backup files.



4. Shared files:





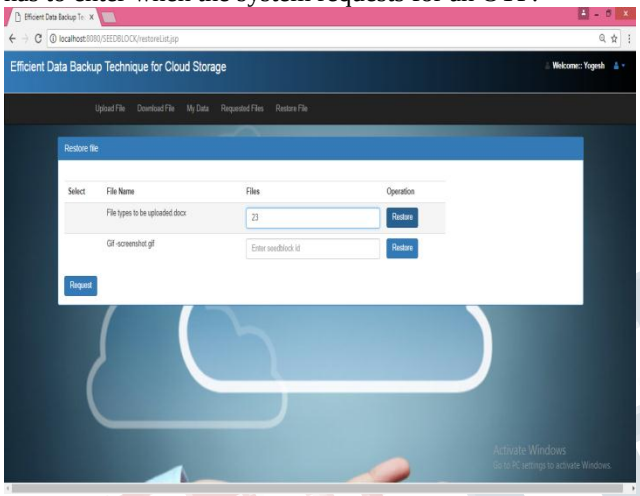
All the files that are shared with all the users publicly, will be displayed, all the files that are shared publicly will have Read-Only permission to all the users

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 3, March 2018

5. Cloud Admin:
The cloud admin has the authority to view all the users in the system. By selecting a particular user the admin can view only the attributes of the files that the specific user has stored in the drive or has recovered them. He will only be able to view the file id, name and the user's seed id.

6. Recover deleted files:
The deleted files always have a backup in the system. If the user wishes to recover a deleted file he can click the Download button. A key (OTP) will be sent to his email id as well on the registered mobile phone number which he has to enter when the system requests for an OTP.



E. Advantages of proposed system
As the data stored are never forfeited i.e. it's reliable
It recovers data from remote backup server same size of original stored at main cloud.
Low data recovery processing cost which provides benefit of backup, recovery assistances to the greatest number of clients
It gives Privacy since no unauthorized user or third party can retrieve the data.
Less amount of time is needed to the authorized user for data access
Maintenance of cloud computing applications is easier
Integrity and confidentiality of user data is preserveed.

IV. FUTURE SCOPE

The software discussed above is a humble effort to bring more effectiveness to the whole system of recovery of files on cloud. Yet there is a scope for modification and up gradation in the future. We tend to add additional functionality of assigning a specific amount of free space for a user on the drive and ask the user to pay for additional space, if required. Author of the file will have Read, Write and Execute permissions, by default and he can assign privilege permissions to a specific no. of users as desired.

V. CONCLUSION

Thus, a detailed design of the Efficient Data Backup Technique for Cloud Storage system is presented. Proposed system is robust in helping the users to collect information from any remote location and to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason. Experimentation and result analysis shows that SBA focuses on the security concept for the back-up files stored at remote server, without using any of the existing encryption techniques. The use of AES for encryption and decryption purpose of files ensures a secure backup of data for a long period of time.

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. Vedashree N, Praveen Kumar KC, Anilkumar G, 2015, "Data Recovery in Cloud Environment Using Seed Block Algorithm", IJCSIT, Vol. 6 (5), 2015, 4593-4598.

6. Somesh P. Badhel, Prof. Vikrant Chole, 2014, "A Review on Data Back-up Techniques for Cloud Computing", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.12, December- 2014, pg. 538-542.

7. J.Sangeetha Priya, Asifa Saman, S.Neevedha, V.Suganya, M.J Zainiya Nazrin, 2015, "Data Recovery Technique using Seed Block Algorithm

for Cloud Computing", IJSRD - International Journal for Scientific Research & Development| Vol. 3, Issue 01, 2015 | ISSN (online): 2321-0613.

8. Ms. Kruti Sharma, Prof K. R. Singh, 2013, "Seed Block Algorithm: A Remote Smart Data Back-up Technique for Cloud Computing", 2013 International Conference on Communication Systems and Network Technologies.

9. Somesh P. Badhel, Prof. Vikrant Chole, 2015, "An Efficient and Secure Remote Data Back-up Technique for Cloud Computing", International Journal of Computer Science and Mobile Computing, Vol.4 Issue.6, June- 2015, pg. 361-369.

10. Douglas Selent, 2010, "Advanced Encryption Standard", Rivier Academic Journal, Volume 6, 2010

11. Nentawe Y. Goshwe, "Data Encryption and Decryption Using RSA Algorithm in a Network Environment", IJCSNS International Journal of Computer Science and Network Security, July 2013

12. Usman, Muhammad, and Usman Akram. "Ensuring Data Security by AES for Global Software Development in Cloud Computing." IT Convergence and Security (ICITCS), 2014 International Conference on. IEEE, 2014.

13. Bertoni, Guido, et al. "Error analysis and detection procedures for a hardware implementation of the advanced encryption standard." Computers, IEEE Transactions on 52.4 (2003): 492-505.

14. Gaurav Raj, Ram Charan Kesireddi and Shruti Gupta, "Enhancement of security mechanism for confidential data using AES-128, 192 and 256bit encryption in cloud", 2015 1st International Conference on Next Generation Computing Technologies (NGCT-2015) Dehradun, India, Sept 2015.

AUTHORS' BIOGRAPHY

Prof. Yogesh Gite is currently working as Assistant Professor at Dilkap Research Institute of Engineering & Management Studies (DRIEMS). His areas of interest are Cloud Computing, Network Security, Cloud forensics, Algorithm, Database. He is perusing Master degree in Computer Engineering at ARMIET.

Prof. Ankush Pawar is a Research Scholar in Computer Science and Engineering at Visvesvaraya Technological University, Belagavi, Karnataka. He is currently working as Associate Professor in Computer Engineering Department in Dilkap Research Institute of Engineering and Management Studies, Neral, Maharashtra. His main research areas of interests are Cloud Computing, IoT, Storage and Network Security

Dr.Shashikant Ghumbre is currently working as Head of Department and Professor in Computer Engineering at Government College of Engineering, Avasari. He received his PhD degree in Computer Engineering from Pune University in 2012. His main research areas of interest are Networking, Artificial Intelligence, and Database Administration. He is also interested in security analysis of IoT.