doi:10.5937/jaes16-15732 Paper number: 16(2018)1, 510, 125 - 131

PREVENTING DATA LOSS EVEN WHEN THE SECURITY SYSTEM COMPROMISE

Renu S* Dr.S H Krishna Veni

Noorul Islam University, Kumaracoil, Thuckley, Tamil Nadu, India

Storage as a service will diminish the risk of data storage but it leads to the era of data storage security. Studies show that multi cloud or inter cloud or clouds of cloud storage are eminent techniques which reduce the traditional risk of storage cloud. Replicating all data into numerous clouds will lead to monetary detriment as well as gigantic growth of electronic data. Inorder to overcome this horrible situation we have proposed a clouds tree to store files into multiple locations. Clouds tree is a hierarchical arrangement of public CSPs. There is no interconnection or configuration between CSPs but the relation is depends on file storage. The files which are more sensitive and capable to achieve business continuity are needed to store multiple locations. File splitting, encryption and storing different CSPs are the technique behind this work. Double or triple authentication schemes are provided to access sensitive or private data.

Key words: Clouds tree, Storage management, Data replication, Data security

INTRODUCTION

Low cost and high availability are the key circumstances to magnetize the customers into storage cloud. The facilities of the cloud increase rapidly and the security threads and challenges growing exponentially [06]. Studies and researches are conducted all over the world to overcome these uncertainties. The core domains which researchers concentrated on cloud security are Cloud Computing Architectural Framework, Governance and Enterprise Risk Management, Legal Issues: Contracts & Electronic Discovery, Compliance & Audit Management, Information Management & Data Security, Interoperability & Portability, Traditional Security, Business Continuity & Discovery Recovery, Data Center Operations, Incident Response, Application Security, Encryption & Key Management, Identity, Entitlement & Access Management, Virtualization, and Security as a Service etc.

A few countries introduce new frameworks to standardize the level of protection against threats and vulnerabilities unique to cloud computing. FedRAMP and C5 are the frameworks implemented in US and Germany respectively [01] . Authorization is a major issue while we store our private information into the cloud environment. Mobile based continuous user authentication system is introduced to ensure the protection of cloud service users [03].The cloud service providers also offer different levels of security to user's data. As the cloud users have lost the control over the data, it is necessary to exhibit its security mechanisms to increase the trust and transparency of the cloud consumers. Data auditing and monitoring mechanisms help to detect and demonstrate security flaws [04].The security flaws [14] are always a nightmare to outsource data [15].

Researchers have addressing all the security issues related to data security. Some of the researches include work flow scheduling [09]; secure data sharing [10] , privacy preserving storage and retrieval in multi cloud [11] etc is focusing cloud data security in single and multi cloud environments. We can't make a single framework to focus all issues. The only solution is to manage available technologies in a hasty manner. Here, we have address the major security factors such as confidentiality, integrity and availability.

RELATED WORKS

M. A. Aman and E.K.Cetinkaya in 2017 proposed [02] a cloud security system which addresses the performance of cloud based backup services. This system focuses three dimensional services to the backup data. The services include selection of encryption intensity, safe duplication and querying on encrypted data. The user has a facility to select the strength of encryption standard of their files.

A. F.Barsoum and M.A. Hassa in 2015 proposed [05] a model describing multicopy dynamic data procession. Most of the cloud consumers opted data outsourcing is to ensure the availability and business continuity. The cloud services are using pay –as-you-go model, monetary loss is less compared with the data processing, handling and storing charges. More consumers were demanded to store their data into multiple data centers to increase the level of scalability, durability and availability. The customers have to be trusted to the CSP that the service contract will be followed. The map-based provable multicopy dynamic data possession (MB-PMDDP) scheme ensures by replicating data into number of locations, outsourcing and updating of dynamic data and authorized users can seamlessly access the file copies in different locations.

** Kumaracoil, Thuckalay, Kanyakumari, Tamil Nadu 629180, India, renus23@gmail.com*

Y.Chen et.al in 2015 proposed [07] network coding system to ensure data recovery with high reliability compared to the existing erasure coding and replication methods. As the cloud data centers are located all over the world, it easy to eavesdropped repaired data anywhere in the network from local datacenter to its remote backup site. This process is called link eavesdropping and is clearly mentioned here. The network coded cloud storage system uses a small repair bandwidth and a symmetric design methodology to identify the data recovery system parameters for specific security level.

S.Namasudra and P.Roy in 2017 proposed [08] a cloud security model which focuses data security and access control issues in cloud computing. Access control means that user can avail any type of resources from the server. The key issues addressed to improve the efficiency of data access control models are high searching time for providing public key of the data owner, high data accessing time, maintenance of database, etc.

A. Bessani et.al in 2013 proposed [13] a novel storage system DEPSKY, Dependable and secures storage in a cloud-of-clouds. It is a combination of individual commercial clouds for the purpose of data storage. DEPSKY can be accessed by invoking equivalent operations in a group of individual clouds. It addresses the major limitations of single cloud such as loss of availability, loss and corruption of data [16],loss of privacy and vendor lock-in. DEPSKY is an application of multi cloud which helps to improve the availability, integrity and confidentiality of the data stored in the cloud through encryption, encoding and replication of the data on various clouds that form a cloud-of-clouds. The deployed system has four commercial clouds and used a Planet Lab to support client's usage from different countries. The performance of DEPSKY is very high compared to single cloud service providers but the monetary costs of using DEPSKY is twice the cost compared with single cloud.

M. A. AlZain et.al in 2012 has conducted a paper survey [17] related to the security and performance of single cloud and multi clouds such as cloud-of-clouds. This survey concludes that as the performance and security of the multi cloud is high compared with single cloud but it receives less attention from research community. This paper focused to promote the use of multi-clouds due to its capability to diminish security threats that affect the users.

Bowers et.al in 2009 proposed [18] a protocol named High Availability and Integrity (HAIL) to controls multiple clouds. It is a distributed cryptographic system which allows a set of servers to ensure that the client's stored data is retrievable and integral. HAIL provides a software layer which concentrates on availability and integrity of the stored data in an inter-cloud but it does not offers an assurance of confidentiality.

PROPOSED SYSTEM

The traditional security concerns which are stay alive with different names that are unauthorized disclosure, unauthorized modification, unauthorized access, and data loss. The entire security threats are related to these features with varying proportions. It is not possible to ensure analogous security to all the files which are outsourced from the organization. We have concentrated a security system which ensures protection from data loss even when the security system or the CSPs get compromise. A complete security package needs proper solution for Data classification, Encryption, Data Storage and Data Access. We can use an automated file classifier to classify organizational data into four main categories such as Sensitive, Private, Protected and Public was proposed by Renu.S and Dr.S H Krishnaveni in 2017 [12].The loss of Sensitive and Private Files may cause the existence, reputation, financial loss etc of the organization.

An effective encryption scheme, Key management and integrity check are the necessary requirement of ensuring confidentiality of the file. A combined encryption scheme is effective to encrypt different files with same security. Combined approach means combining different encryption schemes make a crypto structure for each file and map structure with file-ids. As the numbers of sensitive files are less, developing crypto structure to each file is effortless.

An intelligent cryptographic approach may reduce the risk of data replication, insider attack, outside attack, business continuity etc. A dummy file is also encrypted and combines with the encrypted sensitive file blocks.

Let F be a sensitive file with security code 111 and K1 be a key for encryption.

$$C1 = f(E(F,K_1))$$

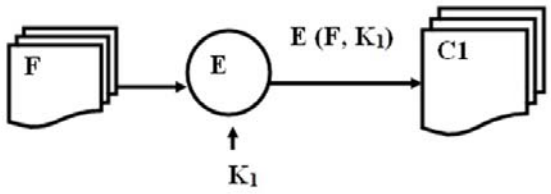

Figure1: Encryption Original File

C1 is the cipher text of original sensitive file. Let D be a dummy file with key K2 and C2 be the cipher text after encryption.

$$C2 = f(E(D, K_1))$$

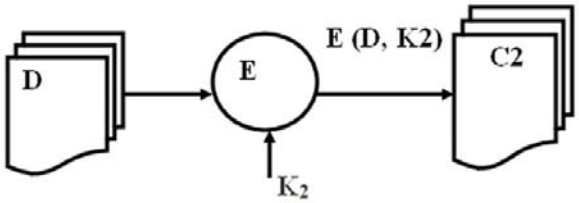

Figure 2: Encryption of Dummy file

The cipher texts C1 and C2 contains number of small blocks,

$$C1 = \{C1_1, C1_2, \ldots C1_n\}$$

$$C2 = \{C2_1, C2_2 \ldots C2_n\}$$

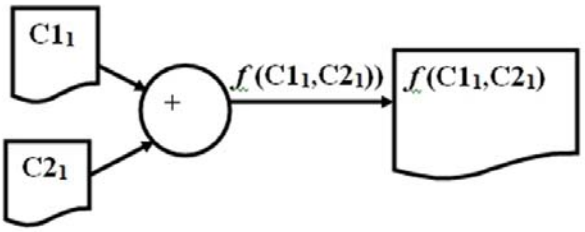

Figure 3: Combining files blocks

$$C = f(C1_1, C2_1)$$

Combine file blocks by using any one of the available function. Create a mapping table which contains file id, security code, keywords for searching the file blocks and user permissions. Let 'key' be a variable used to be searched for file searching. A Hash function is applied to the key value and is appended with the data blocks before sending to the cloud storage.

$$S = f(C+H\,(key))$$

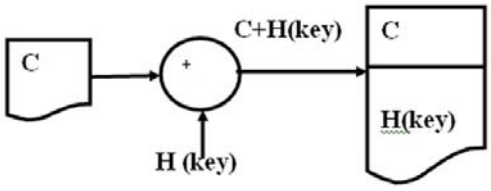

Figure 4: Appending a hash function with Cipher blocks

Let S be a file which contains cipher texts of original file, dummy file and hash value of keywords used for file searching. It will be more secure to send each file block into separate CSPs. Studies show that multi cloud or inter cloud or clouds of cloud storage are eminent techniques which reduce the traditional risk of storage cloud. Storing files into multiple locations will enhance the availability, business continuity, integrity, data backup etc. As the cloud systems are pay –as-you-go model, replicating the entire data into number of storage clouds will leads to monetary loss. Replicating every file into number of clouds is superfluous and futile. Replicating essential and available files is a must to highlight availability and business.

CLOUD TREE

Cloud tree is an arrangement of number of single CSPs in a hierarchical order. Before arranging clouds in a tree we have to execute a detailed study about the security features, pricing, SLA, and all other functions related to storage. The most suitable cloud is set as the root of the tree. Here, we have set a cloud tree with seven individual CSPs.

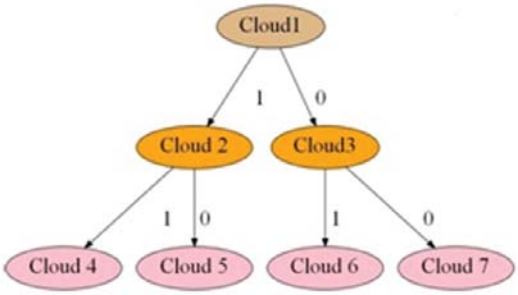

Figure 5: Cloud tree

There is no internal or external configuration with the selected CSPs, but the relationship is resided in storage management. We can use different information management technique to store files into number of locations. Mapping tables keep the information for entire sensitive files. Cloud trees are focused to prevent internal attacks. The arrangements of cloud tree kept secret and use path values or cloud id to store files into public clouds.

Table 1: Example File Mapping table

File id	File blocks	Cloud ids	Encryption details	Hash values
F001	S1,S3,S5	1,3,5	3DES	3a, ab, 4f
F001	S2,S4,S6	2,4,6	3DES	45, b3, a1

FILE SEARCHING

We wouldn't permit file searching outside the organization. A hash value of the file searching keyword is appended with the cipher text blocks. User request is processed and the hash value is send to the corresponding cloud storage. Direct searching is performed at the cloud service provider and the link is enabled to download the corresponding file. A link which includes an encryption application to encrypt the corresponding file is also enabled for the user. We use different combinations of encryption techniques for each file.

ACCESS CONTROL MECHANISMS

Each cloud has different access control mechanisms. CSPs having confidential data have executed compact access control mechanisms. Sensitive and protected data have high confidentiality, a double or triple authentication method is a must to access these types of data. User name and password and file permissions are kept in advance.

A triple authentication method is used to access critical data from the cloud storage. No file searching or key searching is permitted at the cloud. User having permission to access critical data has previously obtains a user name and password for logged in and a separate password for file access. User sends username and password to logging his account. After a successful verification send an OPT to users available mobile number. User resends his access password with OTP for next level authentication. Data manager check the file access permission to the user and after a successful authentication a link is enabled and user can download the file through this link. A single authentication is enough to access file having low confidentiality. Data Manager is a person who has the authority to send and retrieve files from and to the cloud server.

Analysis

Petri net tools are enough to analyze the workflow of file process and file access techniques.

WORKFLOW ANALYSIS

Analysis of the processing of file before outsourcing is analyses with Petri net tools.

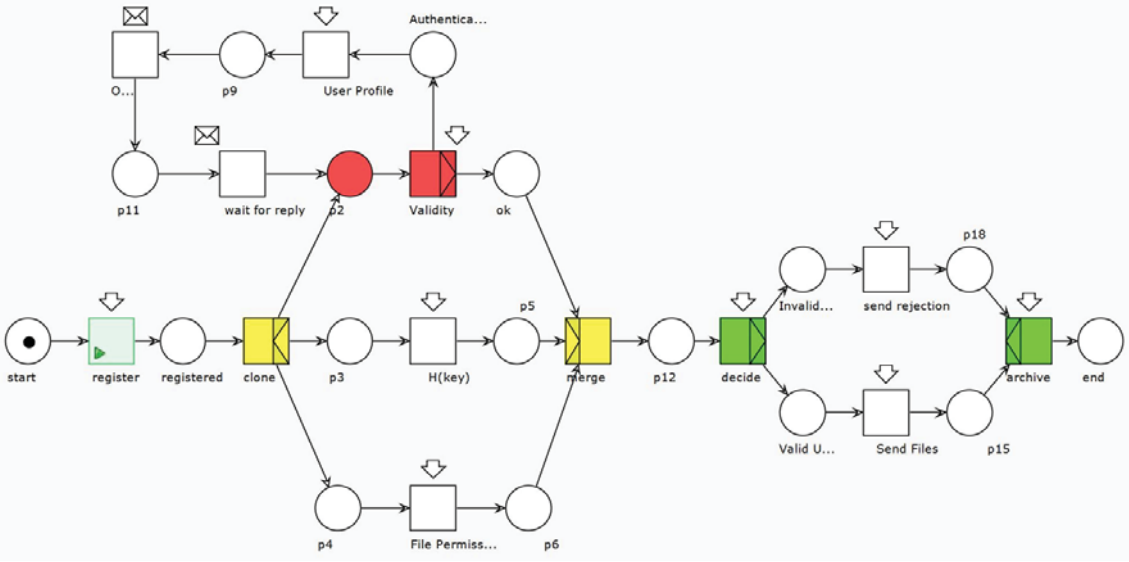

Figure 6: Access control Process with Operator Coloring

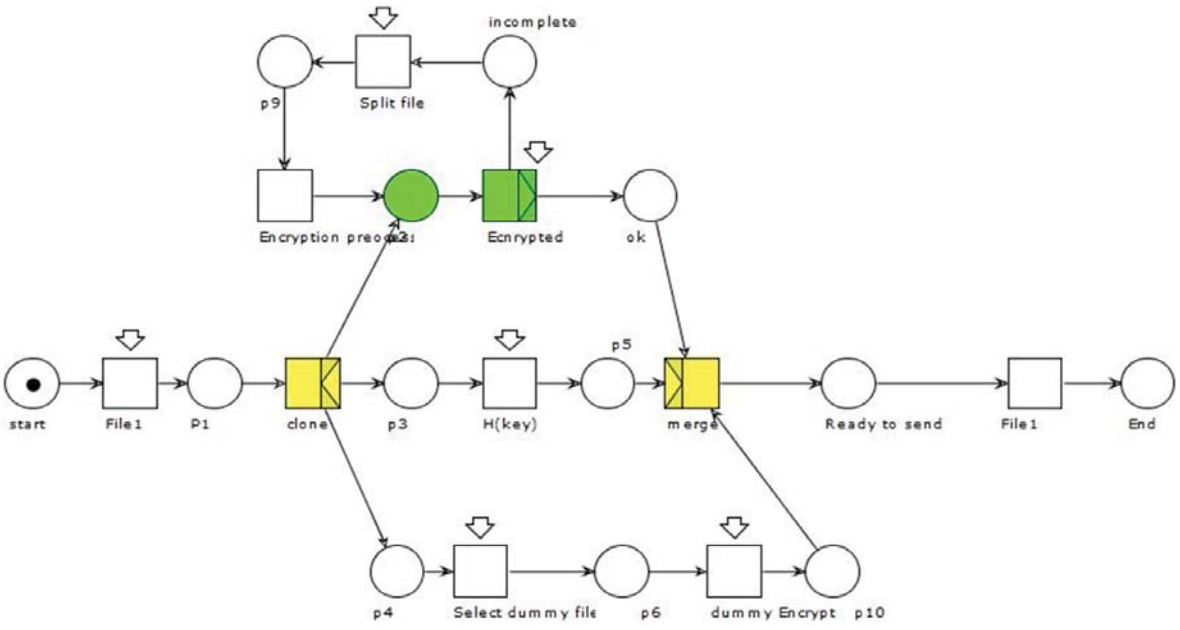

Figure 7: File process with Operator Coloring

PERFORMANCE ANALYSIS

Storing data into multiple clouds will enhance security but it accompanied the threats such as data loss, data replication, and monetary loss etc. Let us analyze how clouds tree prevent data loss, suppose an organization have 500 GB of static data to be outsourced. The available options are single CSP, Multi-clouds and clouds tree. In the case of single CSP complete data is stored in one location; the ownership of the entire confidential data is transferred .It may cause data replication, data loss, malicious insider etc. However in multi-cloud, entire data is replicated into the available CSPs present in the clouds of cloud structure. The security threads still exist. In clouds tree entire data is divided into number of pieces of size 200 GB and store into number of CSPs available in clouds tree, each CSP get small pieces and can solve the replication, data loss, and confidentiality issues. There is not direct splitting to the entire mass, splitting each file systematically. None of the CSPs bear a complete confidential file, but a part of the file.

We have a confidential file of 600 GB , use three different storage technique can explained here.

Case 1. Single CSP: If we prefer to use single CSP, entire file will be stored in a public cloud. Complete file is replicated to the CSP; there is a chance of internal attack, data replication, modification, disclosure etc in a medium level. Data availability also less compared to the other techniques.

Case 2: Multi Cloud: If the user send files to multi cloud, it will automatically replicate into number of CSPs. The probability of internal attack, data replication, modification, and disclosure etc will be high and data availability is extreme compared to the other techniques.

Case 3: Proposed Model: Divide the file into three different pieces and store into different CSP. Each file is divided and store into different CSPs. No one get complete file without the permission of the organization. If any CSPs execute malicious task they can obtain only small pieces of data.

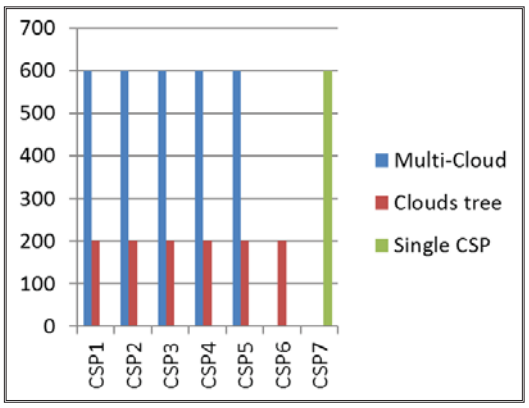

Figure 8: Quantitative analysis - data storage

We can calculate the percentage of data loss if our CSP compromise.

$$\text{Data loss} = \frac{1}{\text{File size}} * \text{size of outsourced file} * 100$$

$$\text{Single CSP, Data Loss\CSP} = \frac{1}{600} * 600 \ * 100 = 100\%$$

$$\text{Multi Cloud, Data Loss\CSP} = \frac{1}{600} * 600 \ * 100 = 100\%$$

$$\text{Proposed, Data loss\CSP} = \frac{1}{600} * 200 \ * 100 = 33.34\%$$

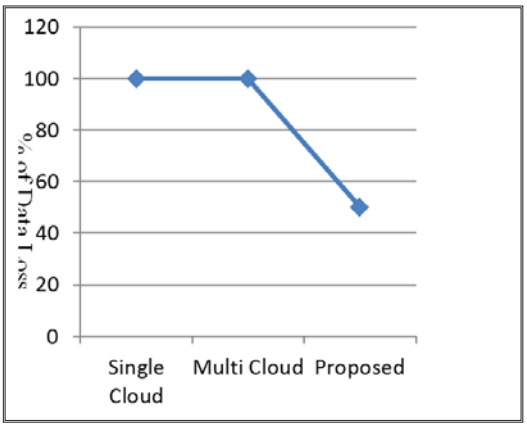

Figure 9: % of data loss even when the CSPs compromise

Let us make a security analysis of a file of size 1GB, single CSP store complete file into one CSP, Multi-cloud store complete file into multiple CSPs, However clouds tree divide file into two equal halves and stored into two different CSPs. Entire storage techniques use strong encryption techniques and access control methods to ensure complete security. However, we will expect the possibility of breaking the security. Proposed system ensures security even when any of the CSPs compromise. Figure 15 shows the rate security attack of different storage system even when CSP compromise.

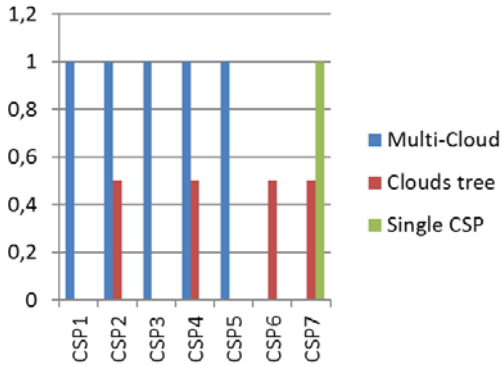

Figure 10: Quantitative analysis-file storage\CSP

The file size and percentage of data attacks are directly proportional. As the file size increases percentage of data attacks also increases. The data attacks include data replication, internal attacks, outside attacks, confidentiality, integrity etc.

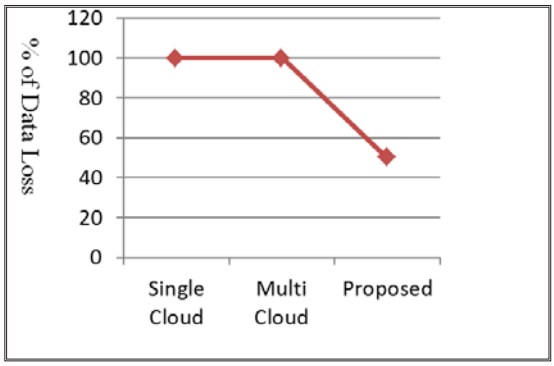

Figure 11: % of data loss even when the CSPs compromise

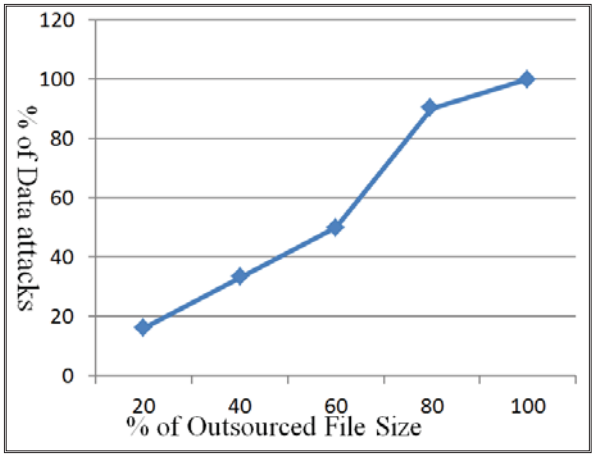

Figure 12: Analysis of data attacks depends on the size of outsourced file

As the file size of the outsourced file increases, data attack\CSP also increases.

DISCUSSIONS

In this section, we can make a functionality analysis of our security system.

- As we are using number of clouds, a through market analysis is possible which helps to identify the functionalities and security services provided by different CSPs.
- Data ownership change is a security threat and can be eliminated by multiple storage facilities, i.e., if any CSP is making ownership problem we can update erroneous data and cancel the agreement with them.
- The possibility of changing the jurisdiction issues can be reduced. Only critical data is stored into multiple locations and which will not cause huge monetary loss.
- As same data is stored into multiple locations, it is easy to check integrity or unauthorized alteration.
- Multiple locations help disaster recoveries which enrich business continuity.
- Only available files are kept in multiple locations which reduces the gigantic growth of e-data.
- Strong encryption technique is executed to data before outsourcing.

No decryption is performed outside the organization.

CONCLUSION

This research shows that we can make a secure data storage environment with available resources. Management of technology is a key concept to ensure data security. Strong encryption and file access technique guarantee security up to a level. The work shows that data attacks and outsourced file size are directly proportional. As the percentage of outsourced file increases, the data attacks also increases. Controlling the size of outsourced file can make a great impact in data security even when the security system compromise.

REFERENCES

1. C.D.Giulio, R.Sprabery, C.Kamhoua, K.Kwit, R.H.Campbell and M.N.Bashir Cloud Standards in Comparison: Are New Security Framework s Improving Cloud Security?" ,Cloud Computing (CLOUD) IEEE 10th international Conference ,Honolulu, USA.,2017.
2. M. A. Aman and E.K.Cetinkaya (2017) Towards Cloud Security Improvement with Encryption Intensity Selection, Design of Reliable Communication Networks, 13 International Conferences, Munich, Germany,2017.
3. Q.Li, L.Wang and T.Kim, Mobile-based continuous user authentication system for cloud security , Network Infrastructure and Digital Content, IEEE International Conference ,Beijing,China,2017.
4. Carlos André Batista de Carvalho, Miguel Franklin De Castro and Rossana Maria De Castro Andrade(2017) Secure Cloud Storage Service for Detection of Security Violations,Cluster, Cloud and Grid computing,17 I EEE /ACM International Conferece,Madrid, Spain, 2017.
5. Ayad F.Barsoum and M.Anwar HAssa,(2015) Provable Multicopy Dynamic Data Possession in Cloud Computing Systems,IEEE Transactions on Information Forensics and Security,Vol. 10,Issue : 3,pp.485-497,2015
6. Sanjay K. Madria, Security and Risk Management in the Cloud ,Computer , Vol. 49,9,pp-110-113,2016.

7. Y.Chen,L.Wang,C.Liao, Eavesdropping Prevention for Network Coding Encrypted Cloud Storage System, IEEE Transactions on Parallel and Distributed Systems,Vol.27,8,pp-2261-2273,2015.
8. S.Namasudra and P.Roy,(2017) Time Saving protocol for data accessing in cloud computing",IET Communications, Vol.11,issue.10, pp.1558-1565,2017
9. H.Chen,X.Zhu, D.Qiu,L.Liu, and Z.Du ,(2017)" Scheduling for Workflows with Security –Sensitive Intermediate Data by Selective Tasks Duplication in Clouds", IEEE Transactions on Parallel and Distributed Systems, Vol. 28,Issue:9, PP. 2674-2688,2017.
10. M.Ali, R.Dhamotharan, E.Khan, S.U.Khan, A.V.Vasilakos, A.Y.Zomaya, SeDaSC:Secure Data Sharing in Clouds, IEEE Systems Journal, Vol.11,issue:2, pp. 395-404,2017.
11. J.Li, D.Lin, A.C.Squicciarini, J.Li,and C.Jia, Towards Privacy-Preserving Storage and Retrieval in Multiple Clouds, IEEE Transactions on Cloud Compputing,Vol.5,Issue: 3,pp.499 -509,2017.
12. Renu.S and Dr.S H Krishnaveni,(2017)An Enhanced CIA tree Using String Matching Algorithms, IJAER, Vol.12, No.16, pp-6123-6126, Sep 2017.
13. A.Bessani, M. Correia, B. Quaresma, F. André and P. Sousa, "DepSky: dependable and secure storage in a cloudof-clouds", ACM Transactions on Storage, Vol. 9, No. 4, Article 12, 2013.
14. Metz, C. DDoS attack rains down on Amazon cloud. The Register. http://www.theregister.co.uk/2009/10/05/amazon bitbucket outage/
15. Raphael, J. The 10 worst cloud outages (and what we can learn from them). Infoworld. http://www. infoworld.com/d/cloud-computing/the-10-worst-cloud-outages-and-what-we-can-learn-them-902,2011.
16. Sarno,(2009) , Microsoft says lost sidekick data will be restored to users. Los Angeles Times.2009.
17. Muhammed A.ALZain , Eric Pardede,Ben Soh, James A Thom,Cloud Computing Secuirty: From Single to Multi-Clouds,45th Hawaii International Conference on System Science,pp- 5490-5499,2012.
18. K.D. Bowers, A. Juels and A. Oprea,HAIL: A high-availability and integrity layer for cloud stor-

Paper submitted: 16.11.2017.
Paper accepted: 08.02.2018.