Volume 5, Issue 8, August 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

# Implementing Data Storage in Cloud Computing with HMAC Encryption Algorithm to Improve Data Security

**Anamika Sirohi, Vishal Shrivastava**
Department of Computer Science and Engineering,
ARYA college of Engineering and IT, Jaipur, Rajasthan, India

***Abstract:*** ***Cloud comes with lots of benefits but still users hesitate to adopt it. Still there are certain issues which are complications in the development of cloud computing. Most protuberant issue prevailing now days is data security at cloud. Main reason or fear in user mind is regarding security whether their data is in insecure hands or is it safe to upload their sensitive data over cloud. To solve this problem of data security a models has been proposed. Our cloud security model plans to keep the most critical data security in cloud computing at different levels like user level, cloud service provider level, third party level and network intruder level. The proposed model provides a way to protect the data, check the integrity and authentication by best possible industry mechanisms. We proposed a model which is extremely secure and is based on data owner model i.e. data is under control of data owner. To maintain data privacy re-encryption is performed with the help of third party and for data integrity Hash Based Message authentication code is generated on encrypted data. Encryption, Clouding, HMAC and Dual substantiation and access management technique has been used which make the proposed model more consistent, scalable, secure and effective to use it in real time applications.***

***Keywords: Cloud computing; encryption; hash mac; data privacy protection, public and private cloud.***

## I. INTRODUCTION

Clouds are essentially large distributed computing facilities that make available their services to third parties on demand. The characterization of a distributed system proposed by Tanenbaum: A distributed system is a collection of independent computers that appears to its users as a single coherent system. Distributed systems often exhibit other properties such as heterogeneity, openness, scalability, transparency, concurrency, continuous availability, and independent failures. To some extent these also characterize clouds, especially in the context of scalability, concurrency, and continuous availability. Three major milestones have led to cloud computing: mainframe computing, cluster computing and grid computing. Mainframes: These were the first examples of large computational facilities leveraging multiple processing units. Mainframes were powerful, highly reliable computers specialized for large data movement and massive input/output (I/O) operations. They were mostly used by large organizations for bulk data processing tasks such as online transactions, enterprise resource planning, and other operations involving the processing of significant amounts of data. Clusters: Cluster computing started as a low-cost alternative to the use of mainframes and supercomputers. The technology advancement that created faster and more powerful mainframes and supercomputers eventually generated an increased availability of cheap commodity machines as a side effect. Starting in the 1980s, clusters become the standard technology for parallel and high-performance computing. Cluster technology contributed considerably to the evolution of tools and frameworks for distributed computing, including Condor, Parallel Virtual Machine (PVM) and Message Passing Interface (MPI). Grids: Grid computing appeared in the early 1990s as an evolution of cluster computing. In an analogy to the power grid, grid computing proposed a new approach to access large computational power, huge storage facilities, and a variety of services. Users can "consume" resources in the same way as they use other utilities such as power, gas, and water. Grids initially developed as aggregations of geographically dispersed clusters by means of Internet connections. Cloud computing is often considered the successor of grid computing. In reality, it embodies aspects of all these three major technologies. Computing clouds are deployed in large datacenters hosted by a single organization that provides services to others. Clouds are characterized by the fact of having virtually infinite capacity, being tolerant to failures, and being always on, as in the case of mainframes. In many cases, the computing nodes that form the infrastructure of computing clouds are commodity machines, as in the case of clusters. The services made available by a cloud vendor are consumed on a pay-per-use basis, and clouds fully implement the utility vision introduced by grid computing.

Cloud computing represents a distributing computing mechanism that by the utilize of the high speed network, data processing is moved from private PC or servers to the remote computer clusters (big data centers owned by the cloud service providers), any user has a potential super computer at hand and can access the data and get the computing capability at any time, from anywhere, you only need to pay for the resources which you have used, don't care about who provide the resources and in what way.

 

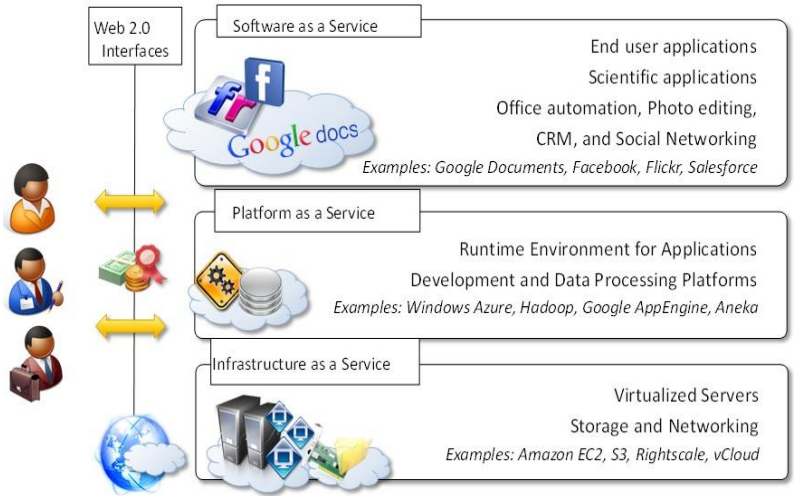


Figure 1: Reference Model of Cloud Computing

Actually, clouds [3] are Internet-based and it tries to disguise complexity for clients. Cloud computing refers to both the applications delivered as services over the Internet and the hardware and software in the datacenters that provide those services. Cloud providers use virtualization technologies combined with self-serviceabilities for computing resources via network infrastructure. In cloud environments, several kinds of virtual machines are hosted on the same physical server as infrastructure. In cloud, costumers must only pay for what they use and have not to pay for local resources which they need to such as storage or infrastructure.

#### HASH MESSAGE AUTHENTICATION CODE (HMAC)

Hash message authentication code (HMAC), a mechanism for message authentication using cryptographic hash functions. HMAC can be used with any iterative FIPS-approved cryptographic hash function, in combination with a shared secret key.The cryptographic strength of HMAC depends on the properties of the underlying hash function. (MSE).

An HMAC function is used by the message sender to produce a value (the MAC) that is formed by condensing the secret key and the message input. The MAC is typically sent to the message receiver along with the message. The receiver computes the MAC on the received message using the same key and HMAC function as was used by the sender, and compares the result computed with the received MAC. If the two values match, the message has been correctly received, and the receiver is assured that the sender is a member of the community of users that share the key.

H MAC uses the following parameters:

B-Block size (in bytes) of the input to the FIPS-approved hash function; e.g., for
SHA-1, B= 64.

H- FIPS-approved hash function, e.g., FIPS 180-1, Secure Hash Algorithm-1 (SHA-1).

Ipad- Inner pad; the byte x'36' repeated B times.

K- Secret key shared between the originator and the intended receiver(s).

K0-The key K with zeros appended to form a B byte key.

L- Block size (in bytes) of the output of the FIPS-approved hash function; for SHA-1,L= 20.

Opad- Outer pad; the byte x'5c' repeated B times.

T- The number of bytes of MAC.

Text- The data on which the HMAC is calculated; the length of the data is n bits, where the maximum value for n depends on the hash algorithm used.

X'N'-Hexadecimal notation, where each 'N' represents 4 binary bits.

||-Concatenation and

Exclusive-Or operation.

## II. LITERATURE SURVEY

In this paper, I have made a review on my topic data security in cloud computing at different levels by reading different kinds of papers and analyzing different techniques which are being used in these papers published by authors which are discussed as follows:

Gupta et.al [1] suggest a scheme where the trust from service provider is not required. The security of data will be in control of the data owner solely. It would mainly contain a tool that would allow the owner of the data to decide about the access rights of his/her data, revocation if any, and notification if any security breaches are in place. This paper also allows a user to search their files in an encrypted database with the help of ranked keyword search which is an improvement over conventional searching techniques.

Lin et.al[2] report a design and implementation of an encrypted cloud storage system that supports multi-user secure indices, allowing efficient search among encrypted documents of multiple users. It also describe a scheme of separating

 

encryption keys and encrypted files and using encrypted (hashed) keywords, neither the plaintext nor the distribution of keywords would be attainable to attackers who only compromise a subset of the servers in the proposed architecture. Experiment results show that keyword search can be performed in real time.

Mowbray et.al [22] gives an overview of issues in privacy protection of personal information in the cloud and describes a variety of approaches that may be used to address these issues. The most appropriate approach varies according to the type of data to be processed or application to be run in the cloud.

Sood et.al [3] proposed approach to ensure data security in cloud computing. In this proposed approach key generation, encryption, indexing of data, user authentication and data integrity is performed by data owner itself. Unfortunately, there will be high overhead on data owner and hence time consuming too. Thilakanathan et.al [4] proposed scheme using proxy re-encryption for security of data. In this scheme data owner encrypt the data using his key piece then proxy encrypt the data using his key piece. Decryption is also carried in similar fashion. However, if proxy is fake then data becomes insecure. Sharma et.al.[5] discussed different service model of cloud computing and highlights the key security issues, challenges and solution at different layers of cloud. Jingwei et.al.[6] discussed efficient model for secure data sharing in cloud. The proposed model consists of user, authority, hybrid cloud and owner. The data is stored at private cloud and data shared is encrypted Encryption technology used is keyword-based encryption. The keys are generated by authority and given to user group for encryption and decryption. The model has some issues like if authority is fake then data is insecure and also it is costly to use the model. Sood et.al [7] proposed the scheme to highly secure the data at cloud. They provided improved data security by using concept of hybrid cloud. In this scheme the sensitive data i.e. about 3%-5% is stored at private cloud and rest of the data at public cloud. This model is applicable to organisations whose sensitive data is about 3%-5%. If the sensitive data increases then this model will prove to be expensive. The white papers [8] of many organisations describes three types of data security models in cloud. First model

Consists of key generation and encryption on data is performed by data owner itself. However this model results in high overhead for data owner. Second model describes encryption performed by data owner and key generation by cloud service provider. Unfortunately, cloud service provider is fake then data is insecure hands. Third model encryption and key generation is control by cloud service provider. If cloud service provider is fake then data is endangered. Hwang et.al [9] proposed business model in which encryption/decryption service and storage as a service of user data were separated i.e. they were not provided by single operator. After encryption/decryption performed system should delete all the data. Varalakshmi et.al [10] proposed system consists of three entities cloud broker, client and cloud storage. Broker handles encryption, hash key, decryption and local database management. According to cloud space available the client files are partitioned into segment and hash values of segments has been generated. When the client needs its file it sends request to broker then broker download the file, partition the file into segments and then calculate the hash values. For checking the data integrity hash values before uploading to the after downloading are matched. If this matches data is un-tampered. Mohamed et.al [11] performed randomness testing on various eight encryption technique namely RC4, RC6, MARS, AES, DES, 3DES, Two-Fish and Blowfish. Xu et.al [12] propose a dynamic user revocation and key refreshing scheme based on cipher policy attribute based encryption technique. In this technique user can be removed anytime without changing keys and also refresh keys without re-encrypting data. Huang et.al [13] proposed scheme that consists of four entities – SSManager, SSGuard, SSCoffer and user. SSGuard do encryption before uploading and uploaded files store at SSCoffer. File encryption key are encrypted by user public key and store at SSManger. For decryption of file QR code is used. User shows QR code to SSGuard to decrypt files. Sur et.al [14] proposed a model in which certificate based Proxy re-encryption scheme is followed before uploading data to cloud. Mowbray et.al [15] gives general overview of protecting data in cloud and describes various approaches to handle this protection. Some of these approaches are available for use now, others are relatively immature, but look promising. The most appropriate approach varies according to the type of data to be processed

Chandel et.al [16] presents a new scheme for secure cloud creation using RC6 (Rivest cipher 6) Encryption algorithm for securing the cloud environment. The results show the performance of proposed technique in public and private cloud. Fan et.al [17] describes Predicate encryption is a novel cryptographic primitive that provides fine-grained control over the accesses to encrypted data. It is often used in secure cloud storage and biometric matching. In this manuscript, we first propose a variant of symmetric predicate encryption, which provides controllable privacy preserving search functionalities, including revocable delegated search and un-decrypt able delegated search. Due to these functionalities, the owner of cloud storage can easily control the lifetimes and search privileges of cloud data. Hashizume et.al [18] describes the security risk associated with services are often outsourced to a third party, which makes it harder to maintain data security and privacy, support data and service availability, and demonstrate compliance. It also discusses security issues as well as to identify and relate vulnerabilities and threats with possible solutions. Lenkala et.al [19] present a risk assessment framework to study the security risk of the cloud carrier between cloud users and two cloud providers. The risk assessment framework leverages the National Vulnerability Database (NVD) to examine the security vulnerabilities of operating systems of routers within the cloud carrier. This framework provides the quantifiable security metrics of each cloud carrier, which enables cloud users to select quality of security services among cloud providers. Such security metric information is very useful in the Service Level Agreement (SLA) negotiation between a cloud user and a cloud provider. It can be also used to build a tool for verifying the commitment of an SLA. Furthermore, implement this framework on Amazon Web Services and Windows Azure, respectively. Our experiments show that the security risks of cloud carriers on these two commercial clouds are significantly different. This finding provides guidance for a network provider to improve the security of cloud carriers.

 

## III. RESEARCH METHODOLOGY

Data security is major issue prevailing in world of cloud computing and to overcome that issue the model has been proposed. Proposed model has been organized in such manner that it give throughout data security in cloud computing at different levels. The different threat levels are: user level, cloud service provider level, third party level and network intruder level. Data is protected against all level of threats. In proposed model data remains private during transit as well as data at rest and from untrusted parities. The goals of proposed model are to provide:

Security at User Level: Data remain secure from dishonest employee of organization or intruder.

Security at cloud service provider: Data remain private from untrusted or fake cloud provider.

Security at Third Party level: Data security against untrustworthy third party if involved data protection.

Network Intruder: Data remains secure during transit or over network form intruders.

Data Confidentiality: Data secrecy is maintained throughout the model i.e. data at rest or over network or during transit.

Data Privacy: Data Leakage is not there without authentication so data remains private.

Overhead: As model is based on owner centric approach so all the overhead will on data owner but this model has been proposed such that data owner overhead should be less.

Data Integrity: To check data tampering over the network by the network intruder during transit of data is kept main concern in model.

Proposed Model has been designed in a way that biggest issue of data security in cloud computing has been resolved and user fearlessly adopt cloud. By comparing with rest of the models it has been concluded that the proposed model is highly secure, data remains private from untrusted parities and free from internal and as well as external threats.The model is divided into two phases and consists of data owner, third party, cloud service provider and user. The phases are:

1. Phase I (Uploading or Data Storage)
2. Phase II (Downloading or Data Retrieval)

**Message Authentication Code (MAC) Generation**

Step1: Message Authentication Code (MAC) is generated on encrypted data using Message-Digest algorithm. Openssl dgst –md5 filename

```
[dataowner@server20 .ssh]$ openssl dgst -md5 file2
MD5(file2)= 0064c057046f128e32e0151e41c787a3
[dataowner@server20 .ssh]$ _
```

Figure2: Message Authentication Code (MAC) Generation

Step2: Encrypt the md5 output same as done before in data encryption

openssl rsautl –encrypt –pubin –inkey PUBLIC_KEY.pem –in md5 –out md

```
[dataowner@server20 .ssh]$ ls
file1  file2  md5  PUBLIC_KEY.pem
[dataowner@server20 .ssh]$ openssl rsautl -encrypt -pubin -inkey PUBLIC_KEY.pem
-in md5 -out md
[dataowner@server20 .ssh]$ ls
file1  file2  md  md5  PUBLIC_KEY.pem
[dataowner@server20 .ssh]$ cat md
x$P^a▪ct▪X9jT+▪▪_Ge▪
▪▪2C▪▪sh▪q▪▪▪▪at▪▪▪IUĎ▪▪ ▪S     ▪~▪▪▪0▪[▪=¦R▪1▪l▪}▪dF▪▪O@▪)B▪▪▪▪▪/▪u}▪▪>I▪▪{~▪▪F
▪▪▪/▪!▪(▪[dataowner@server20 .ssh]$
[dataowner@server20 .ssh]$ _
```

Figure3: Message Authentication Code (MAC) Encryption

**Role-Based User Dual Authentication**

Dual Authentication of user is carried by third party first and further by data owner.

Step1: User Authentication by Third party Data owner has already given list of valid users and third party created database of valid users. User Login to third party using ssh. Third party check its database if valid user then user successfully login.

```
Red Hat Enterprise Linux Server release 6.3 (Santiago)
Kernel 2.6.32-279.el6.x86_64 on an x86_64

server4 login: anamika
Password:
[anamika@server4 ~]$ ssh anamkia@192.168.2.6
The authenticity of host '192.168.2.6 (192.168.2.6)' can't be established.
RSA key fingerprint is 53:98:79:a6:0b:2f:8d:92:2c:07:89:da:4e:91:2e:bd.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '192.168.2.6' (RSA) to the list of known hosts.
anamkia@192.168.2.6's password:
[root@server3 ~]# _
```

Figure4: User Verification by Third party

 

Step2: User send private key without passcode using secure copy

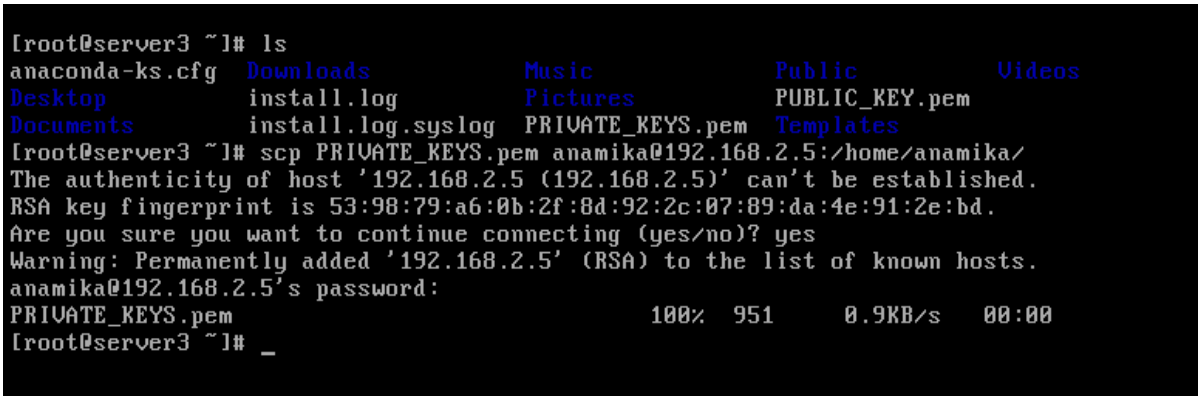


Figure5: Transfer of private key to User by Third party

Step3: User receives private key


Figure6: User receive Private Key

Step4: Third party notifies the data-owner about user and further authentication carried by data owner. Data owner verify the user using smartcard. Data-owner authenticate user without any passcode i.e. through smartcard using ssh using RSA algorithm.

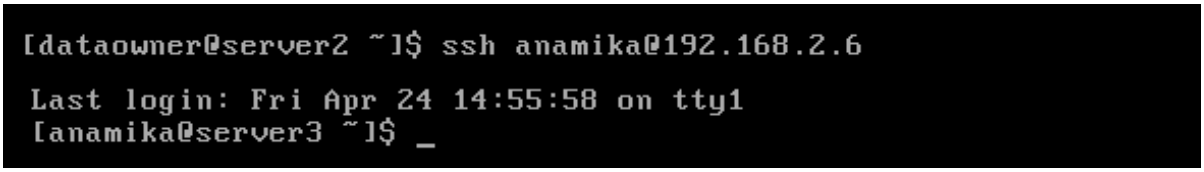


Figure7: User Authentication by Data owner

Step5: Data owner create the file in which login detail of cloud, passcode of private key, detail of MAC and URL of cloud service provider is there. Now data-owner encrypt the logindetail file and transfer encrypted file


Figure8: Creation and Encryption of login detail

Step6: Encrypted file is transferred over network to required user through secure copy (scp)

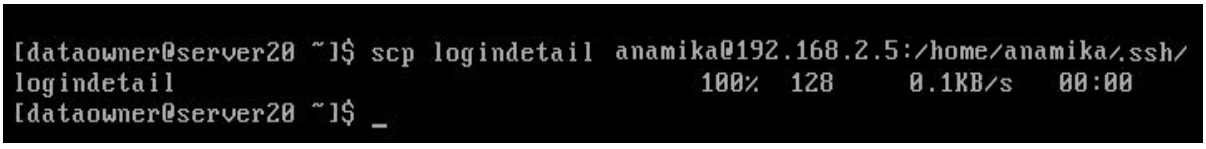


Figure9: Transferred of encrypted login detail to User

 

Step7: Now user decrypt the login detail file using its smartcard


Figure10: Decryption of login detail using smartcard by user

Step8: The login detail file has been decrypted.
Step9: User login to cloud with login id and password as provided by data owner.

## IV. RESULTS ANALYSIS

Proposed model has been organized so that it give throughout data security in cloud computing at different levels. Comparative Analysis between proposed model and other exiting security model has been illustrated below:

Table 6.1: Comparative Analysis

| Parameters | Jing et al [6] | Sood et al [11] | Thilakanthan et al [8] | Li et al [10] | White papers [12] | Proposed model |
|---|---|---|---|---|---|---|
| Confidentiality | Yes | Yes | No | Yes | Yes | Yes |
| Overhead | No | No | No | No | May be | No |
| Authorization | No | Yes | No | No | May be | Yes |
| Encryption | Yes | Yes | Yes | Yes | May be | Yes |
| Dual Verification | No | Yes | No | No | No | Yes |
| Cost Effective | Yes | No | Yes | No | Yes | Yes |

The different levels are: user level, cloud service provider level, third party level and network intruder level. Data is remains private against all level. This model is data owner centric with least overhead and highly secure to adopt in real life while storing and retrieving data from cloud. This model is designed in way to protect data from every aspect. Security analysis of model has been performed at different levels .The security analysis are:

Security at User Level: Role based Dual verification of user is carried out in proposed model. Data is protected against unauthorized access. When user needs data it has to undergo dual verification from third party as well as data owner itself. Once the user verified by both the parties, authorized user can access the data by login to cloud. Access to data is role based i.e. authorized user can read, update or delete data according to data owner wish. Data owner sets the permissions on data for their user. Data is protected from unauthorized access.

Security at Cloud Service Provider: Encrypted data is uploaded over cloud in order to protect data against cloud service provider. Thus even if cloud service provider is fake data is secured.

Security at Third Party level: In proposed model third party act as Key Management Infrastructure. At this level data is verified to be secured against third party. In proposed model it is assumed that third party do not know about cloud service provider. Even if third party knows about cloud service provider, login id and password of authorized user then also third party cannot get cloud login id and password. During user verification by data owner third party will need smart card that was given to authorized user by organization for user authentication. So data is secured even if third party is untrusted.

## IV. CONCLUSION & FUTURE SCOPE

Recent Trends of computing shows cloud computing has made incredible development from past few years. It has brought remarkable change in world of computing. Main reason or fear in user mind is regarding security whether their data is in insecure hands or is it safe to upload their sensitive data over cloud. To solve this problem of data security a models has been proposed. The proposed model provides a way to protect the data, check the integrity and authentication by following the best possible industry mechanisms. In these model data security is checked where ever there is possibility of data threat. Data security is checked at different levels that are user level, cloud service provider level, network intruder level as well as at cloud service provider level. Proposed model is data owner centric with least overhead and highly secure to adopt in real life while storing and retrieving data from cloud. The outcome of this proposed model is that data is kept secure at cloud and feel free without any fear of misuse data can be uploaded. This model provides data confidentiality, rapid availability on demand, data Integrity and minimum overhead to data owner, cost effective and efficient.

 

**REFERENCES**

[1] Sarika Gupta, Sangita Rani Satapathy, Piyush Mehta , Anupam Tripathy, "A Secure and Searchable Data Storage in Cloud Computing", 3rd International Advance Computing Conference (IACC),IEEE,2012

[2] Mao-Pang Lin, Wei-Chih Hong, Chih-Hung Chen, Chen-Mou Cheng, "Design and Implementation of Multi-user Secure Indices for Encrypted Cloud Storage", 11th Annual Conference on Privacy, Security and Trust (PST), IEEE,2013.

[3] Sandeep K. Sood, "*A Combined Approach to Ensure Data Security in Cloud Computing*", Submitted to Journal of Network and Computer Applications, Elsevier Ltd, 2012.

[4] Danan Thilakanatha, Shiping Chen,Surya Nepal, Rafael A. Calvo and Leila Alem, "*A platform for secure monitoring and sharing of generic health data in the Cloud"*, Elsevier Ltd, 2013.

[5] Pardeep Sharma, Sandeep K. Sood, Sumeet Kaur, "*Cloud Implementation Issues and What to Compute on Cloud*", International Journal of Advances in Computer Networks and its Security, vol.1, no. 1, pp. 130-135, 2011.

[6] Jingwei Li, Jin Li, Zheli Liu and Chunfu Jia "*Enabling efficient and secure data sharing in cloud computing*" Concurrency Computat.:Pract Exper.,John Wiley & Sons, Ltd.,2013.

[7] Sandeep K. Sood, "*A Highly Secure Hybrid Security model for Data Security at Cloud* ", Submitted to Security and Communication Networks, John Wiley and Sons (Interscience), Special Issue on Trust and Security in Cloud Computing, 2012.

[8] Amazon Web Services.: "*Encrypting Data at Rest in AWS*", https://aws.amazon.com/whitepapers.

[9] Jing-Jang Hwang, Hung-Kai Chuang, Yi-Chang Hsu and Chien-Hsing Wu, "*A Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service"*, National Science Council of Taiwan Government.

[10] P.Varalakshmi and Hamsavardhini Deventhiran, "*Integrity Checking for Cloud Environment Using Encryption Algorithm"*, IEEE, 2012.

[11] Eman M.Mohamed and Sherif EI-Etriby, "*Randomness Testing of Modem Encryption Techniques in Cloud Environment*", 8th International Conference on Informatics and Systems, 2012.

[12] Zhiqian Xu and Keith M. Martin, "*Dynamic User Revocation and Key Refreshing for Attribute-Based Encryption in Cloud Storage*", International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2012.

[13] Kuan-Ying Huang, Guo-Heng Luo and Shyan-Ming Yuan, "*SSTreasury+: A Secure and* E*lastic Cloud Data Encryption System*", International Conference on Genetic and Evolutionary Computing, IEEE(2012).

[14] Chul Sur, Youngho Park, Sang Uk Shin, Changho Seo and Kyung Hyune Rhee, "Certificate-Based Proxy Re-Encryption for Public Cloud Storage", International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IEEE, 2013.

[15] NarendraChandel, Sanjay Mishra, Neetesh Gupta and AmitSinhal, "Creation of Secure Cloud Environment using RC6", IEEE,2013.

[16] Miranda Mowbray and Siani Pearson, "Protecting Personal Information in Cloud Computing", Springer Verlag, 2012

[17] Chun-I Fan and Shi-Yuan Huang, "Controllable Privacy Preserving Search Based on Symmetric Predicate Encryption in Cloud Storage", International Conference on Cyber Enabled Distributed Computing and Knowledge Discovery, IEEE, 2011.

 