Ensuring Data Integrity Using Cooperative Provable Data Possession for Multi Cloud Environment

Manorama M. Gadge, Rakesh Rajani
Computer Department, Alard College of Engineering And Management, Marunji, Pune, Maharashtra, India

ABSTRACT

Cloud computing is a new and fast growing technology that offers an innovative, efficient and scalable business model for organizations. Cloud computing is type of computing where resources are shared rather than having local servers or personal devices to handle various applications. Cloud computing is a type of internet based computing where different services like server, storage, application, and network are delivered to the organizations. Although almost every business organization is going to adopt this technology due to its various advantages. When cloud computing become more perfect organizations and user will store enormous amount of data on remote cloud storage to achieve remote access, reduced cost, data collection and sharing of other services. To access services from cloud user can have account associated with single and multiple cloud service providers (SPs). Important aspect for cloud computing environment is to maintain integrity of stored data. Implementation of encryption of the information is done in such a way that it will be impossible for the attackers to read the resources sent on the web. Advanced Encryption Standard (AES) and Elliptic Curve Cryptography (ECC) are the methods used for the encryption. Result will be text (cipher) which is decrypted on the receiver's side. AES and ECC algorithm implemented together to provide sound security.

Keywords: Cloud Computing, Cooperative provable data possession, Data Integrity, Multi Cloud

I. INTRODUCTION

Cloud computing has become a faster profit growth point in recent years by providing a comparably low-cost, scalable, position - independent platform for clients' data. In order to lessen the heavy burden of local data storage and maintenance, data is being outsourced into the Cloud. Although storing data into the cloud brings some charming benefits, it also raises some challenging security issues.

The cloud service provider (CSP) offered its customers with kind of services and tools, which are:

a. Software as a Service (SaaS): involves using their cloud infrastructure and cloud platforms to provide customers with software applications. In this service, the user can take advantage of all applications. The end user applications are accessed by users through a web browser, such as Microsoft SharePoint Online. The need for the user to install or maintain additional software is eliminated.

b. Platform as a Service (PasS): enables customers to use the cloud infrastructure; as a service plus operating systems and server applications such as web servers. The user can control the development of web applications and other software and which use a range of programming languages and tools that are supported by the service provider

c. Infrastructure as a Service (IaaS): the registered user may access to physical computing hardware; including CPU, memory, data storage and network connectivity of the service provider. IaaS enables multiple customers referred to as "multiple tenants" using virtualization software. The user gains greater flexibility in access to basic infrastructure.

d. Security as a service (SecaaS): categorize the different types of Security as a Service and to provide guidance to organizations on reasonable implementation practices.

Cloud storage is visualized pools where data and applications are stored which are hosted by the third party. Company, who desires to store their data in the cloud, buy or lease storage capacity from them and use it for their storage needs. Some of the cloud storage benefits are reduce costs, provide more flexibility, reduce IT management of hardware and data, reduce management of web applications through automated updates, and provide greater storage capacity. In spite of these benefits, "cloud" lack in some of the issues like data integrity, data loss, unauthorized access, privacy etc. Data Integrity is very important among the other cloud storage issues. Because data integrity ensured that data is of high quality, correct, consistent and accessible. After moving the data to the cloud, owner hopes that their data and applications are in secured manner. But that hope may fail sometimes the owner's data may be altered or deleted.

In order to ensure the integrity and availability of data in Cloud and enforce the quality of cloud storage service, efficient methods that enable on-demand data correctness verification on behalf of cloud users have to be designed. However, the fact that users no longer have physical possession of data in the cloud prohibits the direct adoption of traditional cryptographic primitives for the purpose of data integrity protection.

Hence, the verification of cloud storage correctness must be conducted without explicit knowledge of the entire data files. The data stored in the cloud may not only be accessed but also be frequently updated by the users, including insertion, deletion, modification, appending, etc. Thus, it is also imperative to support the integration of this dynamic feature into the cloud storage correctness assurance, which makes the system design even more challenging.

II. METHODS AND MATERIAL

A. Literature Survey

The data integrity problem is solved by many proposed systems. All of them fall into public auditability and private auditability. Private auditability provides higher scheme efficiency while public verifiability allows any one, not just the client to challenge the cloud server for correctness of the data storage while keeping no private information. Most of the works for storage security in cloud computing concerned with integrity checking of remotely stored data.

NarnYih Lee et al. [11], propose PDP scheme based on symmetric and non-symmetric key encryption. The proposed scheme is efficient because it does not require more outsourced data to be encrypted and no additional posts on the symbol block also more secure because data is encrypted to prevent unauthorized parties to know its contents. This research also focuses on public verifiability for hybrid cloud where anyone other than owner verify correctness of data stored on the server.

Qian Wang et al. [6], proposes a protocol to obtain efficient data dynamics by improving the existing proof of storage models by manipulating classic Merkle Hash Tree construction for block tag authentication. The paper supports public auditability for correctness of stored data where not just client but anyone allows verifying correctness of data on demand. In public auditability Third party auditor which verifies integrity of dynamic data on behalf of cloud client. The introduction of Third party verifier helps to achieve economics of scale for cloud computing. The research supports data dynamics through block modification, insertion, and deletion since cloud computing services are not limited to only backup of data. This paper explores the problem of integrity verification of outsourced data by doing data dynamic and public auditability simultaneously. The protocol supports batch auditing where multiple auditing tasks from different clients are aggregate into single audit task. For batch auditing bilinear aggregate signature scheme is used in which multiple signatures by distinct user on distinct message are aggregate into a single short signature. This scheme greatly reduces communication cost and provides efficient verification for all messages.

Yan Zhu et al. [3] focus on irretrievable loss of data from cloud due to lack of integrity verification mechanism for distributed data outsourcing. Paper introduces construction of Collaborative Provable Data Possession scheme for hybrid clouds. Hybrid cloud is the cloud computing environment in which an organization provides and manages some resources in-house and has other resources provided externally. Hybrid cloud allow to take advantage of scalability and data migration to support these features paper propose a effective construction of collaborative provable data possession based on homomorphic verifiable responses and hash index hierarchy. Security of proposed scheme is based on multi-prover zero-knowledge proof system which satisfies properties of completeness, knowledge soundness and zero- knowledge. Paper provides effective construction of CPDP provides security against data leakage attack and tag forging attacks. Paper proposes collaborative provable data possession mechanism to support dynamic scalability on multiple storage servers by allowing transparent property to clients. So clients can store and manage the resources in hybrid cloud by incurring a small amount of communication overhead. Before construction of cooperative provable data possession for multi cloud Yan Zhu [4] proposes CPDP for hybrid cloud. Paper focuses on publicly verifiable provable data possession to support privacy protection and dynamic scalability in which client can dynamically access and update their data in hybrid cloud. Paper provides effective construction of cooperative provable data possession along with homomorphic verifiable response and hash index hierarchy. Using PDP clients can verifies availability and integrity of data stored in multiple cloud service providers without any knowledge of where data is geographically located. Paper introduces effective CPDP construction which lessens communication complexity but requires small and constant amount of overhead.

Afterwards Yan Zhu et al. [1] focuses on cooperative provable data possession scheme for integrity verification based on homomorphic verifiable response and hash index hierarchy. Research proves security of the scheme based on multi- prover zero knowledge proof system. This scheme verifies integrity of outsourced data with lower communication and computation overheads as compared to non cooperative approach. While checking integrity for large files, integrity is affected by bilinear mapping due to its high complexity. One of the challenging jobs of the paper is generation of tags with the length irrelevant to size of data.

B. System Architechture

In multi cloud architecture, a data storage service involves three different entities:
Client is an entity either individual user or organization which uses cloud to store large amount of data and depend on cloud for maintenance and computation of data. Cloud service providers (CSPs) who have significant storage and computation resources to manage client data and offer storage services to the client. In multi cloud service providers organize their resources and provide storage service to the client. Trusted Third Party is a trusted entity to which verification parameters are stored and for this verification parameter it proposes

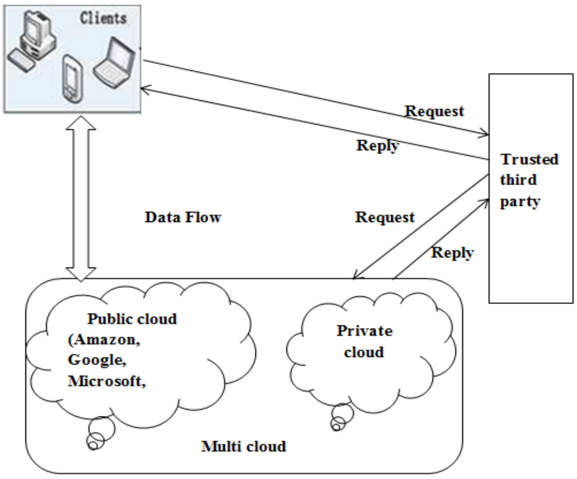


Figure 1: System Architecture for Proposed Model

C. Proposed Algorithm

a. Elliptic Curve Diffie-Hellman Algorithm

Elliptic curve Diffie–Hellman (ECDH) is an anonymous key agreement protocol that allows two parties, each having an elliptic curve public–private key pair, to establish a shared secret over an insecure channel.

i. A particular rational base point P is published in a public domain for use with a particular elliptic curve E(Fq) also published in a public domain.
ii. Alice and Bob choose random integers kA and kB respectively, which they use as private keys.
iii. Alice computes kA*P, Bob computes kB*P and they exchange these values over an insecure network.
iv. Using the information they received from each other and their private keys, both Alice and Bob compute (kA*kB)*P = kA*(kB*P) = kB*(kA*P).
v. This value is then the shared secret that only Alice and Bob possess.

b. AES Algorithm

Rijndael was designed to have the following characteristics:

- Resistance against all known attacks.

- Speed and code compactness on a wide range of platforms.
- Design Simplicity.

Inner Workings of a Round

The algorithm begins with an Add round key stage followed by 9 rounds of four stages and a tenth round of three stages. This applies for both encryption and decryption with the exception that each stage of a round the decryption algorithm is the inverse of its counterpart in the encryption algorithm. The four stages are as follows:

1. Substitute bytes
2. Shift rows
3. Mix Columns
4. Add Round Key

The tenth round simply leaves out the Mix Columns stage. The first nine rounds of the decryption algorithm consist of the following:

1. Inverse Shift rows
2. Inverse Substitute bytes
3. Inverse Add Round Key
4. Inverse Mix Columns

Again, the tenth round simply leaves out the Inverse Mix Columns stage.

Substitute Bytes: This stage (known as SubBytes) is simply a table lookup using a 16×16 matrix of byte values called an s-box. This matrix consists of all the possible combinations of an 8 bit sequence (28 = 16 × 16 = 256). However, the s-box is not just a random permutation of these values and there is a well-defined method for creating the s-box tables.

Shift Row Transformation: This stage is a simple permutation and nothing more. It works as follow:

- The first row of state is not altered.
- The second row is shifted 1 bytes to the left in a circular manner.
- The third row is shifted 2 bytes to the left in a circular manner.
- The fourth row is shifted 3 bytes to the left in a circular manner.

The Inverse Shift Rows transformation (known as InvShiftRows) performs these circular shifts in the opposite direction for each of the last three rows.

Mix Column Transformation: This stage (known as MixColumn) is basically a substitution where each column is operated on individually. Each byte of a column is mapped into a new value that is a function of all four bytes in the column. The transformation can be determined by the following matrix multiplication on state. Each element of the product matrix is the sum of products of elements of one row and one column.

Add Round Key Transformation: In this stage (known as AddRoundKey) the 128 bits of state are bitwise XORed with the 128 bits of the round key. The operation is viewed as a column wise operation between the 4 bytes of a state column and one word of the round key. This transformation is as simple as possible which helps in efficiency but it also affects every bit of state.

III. MATHEMATICAL MODEL

A Cooperative provable data possession scheme S is a collection of two algorithms and an interactive proof system, S = (KeyGen, TagGen, Proof)

KeyGen (1K): It takes a security parameter k as input, and returns a secrete key sk or a public-secrete key pair (pk, sk).

TagGen (sk, F, P): It takes secrete key sk, file F, and set of cloud service providers P, where P= {Pk} and returns the triples (Ts, Vp, Ta). Where Ts is the secrete of tags, Vp= (u, H) is a set of verification parameters u and an index hierarchy H for F, Ta= {Ta (k)} Pk belongs to P denotes a set of all tags, Ta (k) is a tag of the fraction F (k) of F in Pk.

Proof (P, V): It is a protocol of proof of data possession between the CSPs (P = {Pk}) and a verifier (V), that is

$$\Big(\sum_{\mathrm{Pk}\in \mathrm{P}} \mathrm{Pk}\left(\mathrm{F}^{(\mathrm{k})}, \mathrm{Ta}^{(\mathrm{k})}\right) \leftrightarrow \quad \mathrm{V}\Big)\ (\mathrm{Pk}, \mathrm{Vp}) = \begin{cases} 1 & \text{File F} = \{\mathrm{F}^{(\mathrm{K})}\} \text{ is not changed} \\ 0 & \text{File F} = \{\mathrm{F}^{(\mathrm{K})}\} \text{ is changed} \end{cases}$$

Where each Pk takes as input a file F (k) and a set of tags Ta (k), and a public key pk and a set of public parameters Vp is the common input between P and V. At the end of the protocol, V returns a bit {0|1} denoting false and true where, ΣPk∈ P denotes the collaborative computing in Pk ∈ P. The insignificant way to recognize CPDP is check stored data in each cloud one by one. But this would incurs communication and computation overhead on verifier this lessen advantage of cloud storage. To solve this problem we include organizer (O) which is one of the cloud service provider that directly contacts with verifier as follows:

$$\left(\sum_{Pk\in P} \mathrm{Pk}\left(F^{(k)}, Ta^{(k)}\right) \leftrightarrow O \leftrightarrow V\right)(\mathrm{Pk}, \mathrm{Vp})$$

IV. RESULTS AND DISCUSSION

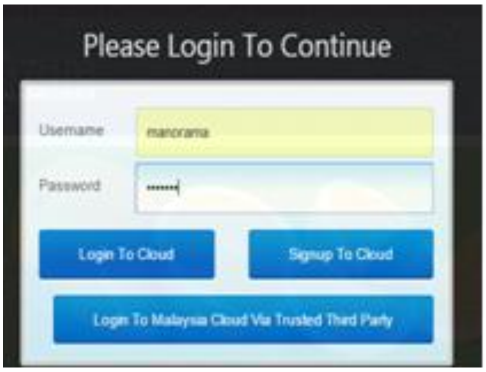
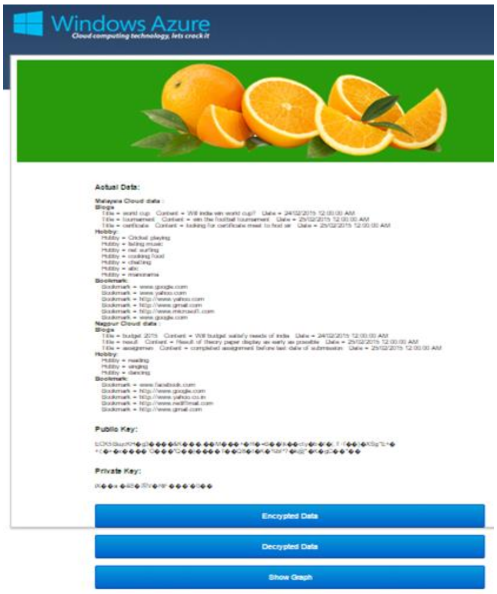


Figure 2: Login to cloud



Actual Data:

Malaysia Cloud data :
Blogs
Title = world cup Content = Will india win world cup? Date = 24/02/2015 12:00:00 AM
Title = tournament Content = win the football tournament Date = 25/02/2015 12:00:00 AM
Title = cerificate Content = looking for certificate meet to hod sir Date = 25/02/2015 12:00:00 AM
Hobby:
Hobby = Cricket playing
Hobby = listing music
Hobby = net surfing
Hobby = cooking food
Hobby = chatting
Hobby = abc
Hobby = manorama
Bookmark:
Bookmark = www.google.com
Bookmark = www.yahoo.com
Bookmark = http://www.yahoo.com
Bookmark = http://www.gmail.com
Bookmark = http://www.microsoft.com
Bookmark = www.google.com
Nagpur Cloud data :
Blogs
Title = budget 2015 Content = Will budget satisfy needs of india Date = 24/02/2015 12:00:00 AM
Title = result Content = Result of theory paper display as early as possible Date = 25/02/2015 12:00:00 AM
Title = assignmen Content = completed assignment before last date of submission Date = 25/02/2015 12:00:00 AM
Hobby:
Hobby = reading
Hobby = singing
Hobby = dancing
Bookmark:
Bookmark = www.facebook.com
Bookmark = http://www.google.com
Bookmark = http://www.yahoo.co.in
Bookmark = http://www.rediffmail.com
Bookmark = http://www.gmail.com

Public Key:

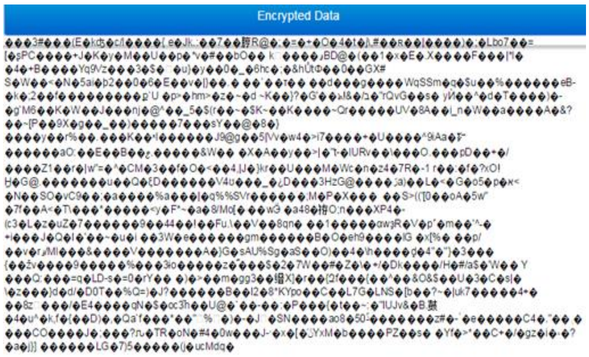
[illegible]

Private Key:

[illegible]

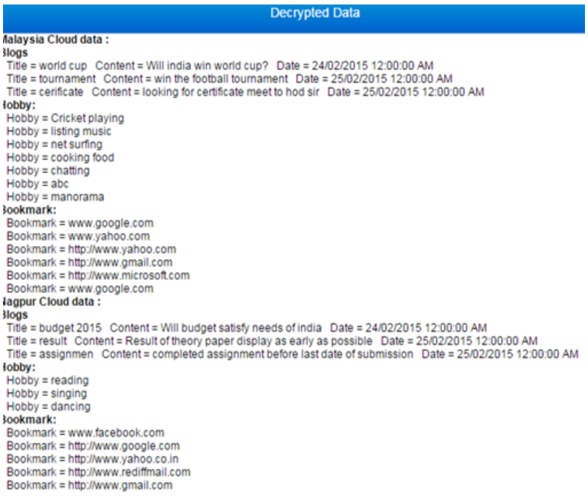
Figure 3: Stored data on both clouds

Above figures shows user can sign up and login to cloud. Data is stored in multi cloud, while retrieving data from multi cloud, clouds communicates with each other by transforming data to one another in encrypted form.



[illegible]

Figure 4: Stored data on Clouds in encrypted form



Malaysia Cloud data :
Blogs
Title = world cup Content = Will india win world cup? Date = 24/02/2015 12:00:00 AM
Title = tournament Content = win the football tournament Date = 25/02/2015 12:00:00 AM
Title = cerificate Content = looking for certificate meet to hod sir Date = 25/02/2015 12:00:00 AM
Hobby:
Hobby = Cricket playing
Hobby = listing music
Hobby = net surfing
Hobby = cooking food
Hobby = chatting
Hobby = abc
Hobby = manorama
Bookmark:
Bookmark = www.google.com
Bookmark = www.yahoo.com
Bookmark = http://www.yahoo.com
Bookmark = http://www.gmail.com
Bookmark = http://www.microsoft.com
Bookmark = www.google.com
Nagpur Cloud data :
Blogs
Title = budget 2015 Content = Will budget satisfy needs of india Date = 24/02/2015 12:00:00 AM
Title = result Content = Result of theory paper display as early as possible Date = 25/02/2015 12:00:00 AM
Title = assignmen Content = completed assignment before last date of submission Date = 25/02/2015 12:00:00 AM
Hobby:
Hobby = reading
Hobby = singing
Hobby = dancing
Bookmark:
Bookmark = www.facebook.com
Bookmark = http://www.google.com
Bookmark = http://www.yahoo.co.in
Bookmark = http://www.rediffmail.com
Bookmark = http://www.gmail.com

Figure 5: Retrieve data from Clouds in decrypted form

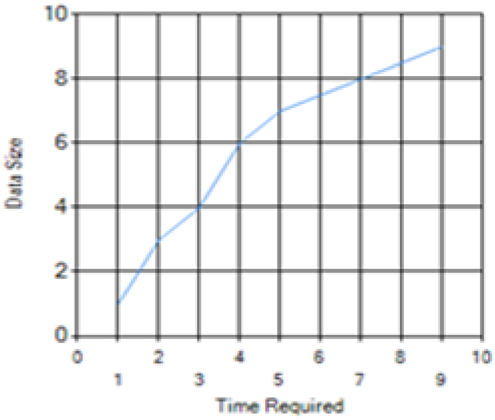


Figure 6: Time required for encrypting data

V. CONCLUSION

We use cryptographic model to provide integrity for client data stored in multi cloud. We focus on exploring way of encryption done; improve some aspects of the algorithm which is already existed and create way for the excellent security. We focus on Cooperative Provable Data Possession to confirm exactness of data in cloud storage. By delivering supervision or control of multi cloud to third party this scheme minimizes amount of computation on both client and server side. The Proposed Elliptic Curve Cryptography system and AES algorithm provides much security in storing data in clouds. The Proposed schemes are efficient which makes minimum use of computation and communication overheads.

VI. REFERENCES

[1] Yan Zhu, Hongxin Hu, Gail-JoonAhn,"Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage" IEEE Transactions On Parallel And Distributed Systems,Digital Object Indentifier 10.1109/TPDS 2012.66 April 2012.

[2] Jing-Jang Hwang, Hung-Kai Chuang, Yi-Chang Hsu, Chien-Hsing Wu, "A Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service" 978-1-4244-9224-4/11/$26.00 ©2011 IEEE

[3] Y. Zhu, H. Hu, G.-J. Ahn, Y. Han, and S. Chen, "Collaborative integrity verification in hybrid clouds," in IEEE Conference on the 7th International Conference on Collaborative Computing: Networking Applications and Worksharing, collaborateCom, rlando, Florida, USA, October 15-18, 2011, pp. 197–206.

[4] Nils Gura, Arun Patel, Arvinderpal Wander, Hans Eberle, and Sheueling Chang Shantz, "Comparing Elliptic Curve Cryptography and RSA on 8-Bit CPUs" International Association for Cryptologic Research 2004, CHES 2004, LNCS 3156, pp. 119–132, 2004.

[5] Yan Zhu, Gail-Joon Ahn, Hongxin Hu, Stephen S. Yau, Ho G. An, Shimin Chen, "Dynamic Audit Services for Outsourced Storages in Clouds" Digital Object Indentifier 10.1109/TSC.2011.51, 1939-1374/11/$26.00 © 2011 IEEE

[6] Kamlesh Gupta, Sanjay Silakari,"ECC over RSA for Asymmetric Encryption: A Review" ISSN (Online): 1694-0814 IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, May 2011

[7] Dalia Attas ,Omar Batrafi, " Efficient Integrity Checking Technique for Securing Client Data in Cloud Computing" International Journal of Electrical & Computer Sciences IJECS-IJENS Vol: 11 No: 05 October 2011

[8] [4] Yan Zhu, Huaixi Wang, Zexing Hu1, Gail- JoonAhn, Hongxin Hu, Stephen S. Yau, "Efficient Provable Data Possession for Hybrid Clouds" CCS'10, October 4–8, 2010, Chicago, Illinois, USA. ACM978-1-4503-0244-9/10/10

[9] D. Sravana Kumar, C H. Suneetha, A. Chandrasekh A R, "Encryption Of Data Using Elliptic Curve Over Finite Fields" International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.1, January 2012

[10] Q. Wang, C. Wang, I. Li, K. Ren, and W. Lou, Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing " IEEE Transactions On Parallel And Distributed Systems, Vol. 22, No. 5, May 2011

[11] Q. Wang, C. Wang, I. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in ESORICS, 2009, pp. 3 55-370.

[12] Kavita Murugesan, shilpa Sudheendran, "Ensuring User Security and Data integrity in Multi- Cloud" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013

[13] Ram Ratan Ahirwal, Manoj Ahke, "Elliptic Curve Diffie-Hellman Key Exchange Algorithm for Securing Hypertext Information on Wide Area Network" International Journal of Computer Science and Information Technologies, Vol. 4 (2) , 2013, 363 - 368

[14] Neha Jha, Brajesh Patel, "Forward Secrecy For Google HTTPS using Elliptic Curve Diffie-Hellman Key Exchange Algorithm" ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 9, November 2012

[15] Narn Yih Lee, Yun Kuan Chang, "Hybrid Provable Data Possession at Untrusted Stores In Cloud Computing," in IEEE Conference on the 17 th International Conference On Parallel And Distributed Systems10.1109 / ICTPDS 2011.70

[16] Chandra Sekhar Golagana, M.Sreedhar, G.Chinna Babu, "A Novel Application for Integrity Verification in Multi-Cloud Storage by using Provable data Possession" International Journal of Application or Innovation in Engineering & Management ISSN 2319 – 4847 Volume 2, Issue 11, November 2013

[17] Saranya Eswaran, Dr.Sunitha Abburu, " Identifying Data Integrity in the Cloud Storage" ISSN (Online): 1694-0814 IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012

[18] M.Venkatesh, M.R.Sumalatha, Mr.C.SelvaKumar "Improving Public Auditability, Data Possession in Data Storage Security for Cloud Computing" ISBN: 978-1-4673-1601-9/12/$31.00 ©2012 IEEE