ENABLING DATA INTEGRITY PROTECTION IN CLOUD STORAGE USING CRYPTOGRAPHY

Mrs. B. Meena Preethi[1], Mr. A. C. Sountharraj[2], Ms. S. Aishwarya[3], Ms. R. Rajeshwari [4]
[1,2]Assistant Professors, Department of BCA & M.Sc. SS
[3,4]V M.Sc. SS Students, Department of BCA & M.Sc. SS
Sri Krishna Arts and Science College, Kuniamuthur, Coimbatore-8

***Abstract*- Cloud computing is a general term for anything that involves delivering hosted services over the Internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Cloud computing promises several attractive benefits for businesses and end users. From users' perspective, including both individuals and enterprises, storing data remotely into the cloud in a flexible on-demand manner brings appealing benefits: relief of the burden for storage management, universal data access with the independent geographical locations, and avoidance of capital expenditure on hardware, software, and personnel maintenances, etc. Cryptography is one of the most important security technologies which used to secure the data transmission and the data itself. As the time and challenge growth, the cryptography also grows up with variety of encryption techniques and algorithms. It protects against unauthorized parties by preventing unauthorized alteration of use. It uses a cryptographic system to transform a plaintext into a cipher text, using most of the time a key. Crypto cloud computing is a new secure cloud computing architecture. It can provide protection of information security at the system level, and allows users access to share services conveniently and accurately. In the crypto cloud computing system, each entity encrypts data using their own private key. While fulfilling their own functions of information exchange and processing, all these elements will use the public key and private key to perform authentication first. In this way, crypto cloud system guarantees the security and credibility of information exchange.**

I. INTRODUCTION

Cloud Computing has been envisioned as the next-generation architecture of IT enterprise, due to its long list of unprecedented advantages in the IT history: on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing and transference of risk. As a disruptive technology with profound implications, Cloud Computing is transforming the very nature of how businesses use information technology. One fundamental aspect of this paradigm shifting is that data is being centralized or outsourced into the Cloud. From users' perspective, including both individuals and enterprises, storing data remotely into the cloud in a flexible on-demand manner brings appealing benefits: relief of the burden for storage management, universal data access with independent geographical locations, and avoidance of capital expenditure on hardware, software, and personnel maintenances, etc . While these advantages of using clouds are unarguable, due to the opaqueness of the Cloud as separate administrative entities, the internal operation details of cloud service providers (CSP) may not be known by cloud users data outsourcing is also relinquishing user's ultimate control over the fate of their data. Cloud storage offers an on-demand data outsourcing service model, and is gaining popularity due to its elasticity and low maintenance cost. However, security concerns arise when data storage is outsourced to third party cloud storage providers. It is desirable to enable cloud clients to verify the integrity of their outsourced data, in case their data have been accidentally corrupted or maliciously compromised by insider/outsider attacks. One major use of cloud storage is long-term archival, which represents a workload that is written once and rarely read. While the stored data are rarely read, it remains necessary to ensure its integrity for disaster recovery or compliance with legal requirements since it is typical to have a huge amount of archived data, whole-file checking becomes prohibitive. The architecture of the Cloud Computing involves multiple cloud components interacting with each

other about the various data they are holding on too, thus helping the user to get to the required data on a faster rate. When it comes to Cloud it's more focused upon the frontend and the back end. The front end is the User who requires the data, whereas the backend is the numerous data storage device, server which makes the Cloud.

II. METHODOLOGY

A. Study On Existing System

The practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer is known as cloud computing. Many of the disadvantages of cloud computing are due to the fact that the technology is still relatively new. In the existing system un-encoded data has been used by FMSR codes which provides less security and allows the third party auditor to become a malicious attacker.

- In case of software failure data mart loss the information of particular client duo to different causes like network down, file outage etc. Data mart takes the copy from backup warehouse. It increases the availability of information. Data mart S1 loss the information P1, then it can take information S1 from backup warehouse and reconfigure it. If data mart S2 and S3 lost the information then they can also able to recover the information from backup ware house.
- The data mart is crashes or down also impact on the availability of information. The purposed system also removes that drawback. If any data mart is crashes or down then client's request also able to extract the data from backup warehouse. In Purposed scenario data mart S1 is fail and not responding the user request. In this case the part of information P1 is lost. The purposed system allow user to extract the information from backup ware house. The availability of data mart also affect on security of information. In case of large no of data marts the data divide in more parts and store different parts in different data marts. Each data marts have very small part of information. If any data mart is hacked by attacker then it can take only small part of information.

Like all new technologies, cloud computing has its fair share of disadvantages. While this affects in general all users, private users or smaller companies are the ones who may feel slightly more disadvantaged than larger businesses such as Network Connections Dependency, Connecting Peripherals, Costs, Data Ownership, No Hard Drive, Security.

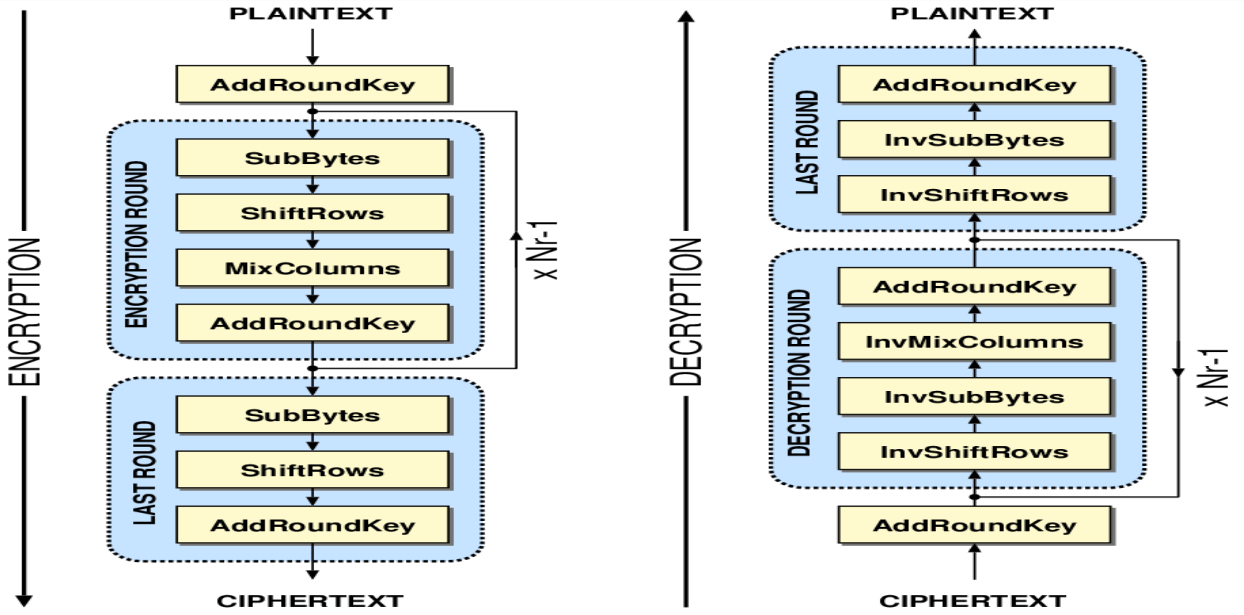

Fig.2 Working of AES

B. Problem Definition

For checking the integrity of the huge amount of archived data the whole-file checking method becomes prohibitive. Proof of irretrievability and proof of data possession are used to verify the integrity of the large data. If the data is stored in the cloud means the data integrity confirmation is very important. It has more chance to modification in the data. Then the developer should repair the corrupted data and restore the original data. Store all the data in a single server is vulnerable to the single point-of-failure problem and vendor lock-ins. The coding has a lower storage overhead than replication under the same fault tolerance level. It results in permanent data loss in the server. These schemes can only provide the detection of corrupted data. Time consumption for find the original data is very high. When it comes to Security, cloud really suffers a lot. There is also a possibility where a malicious user can penetrate the cloud by impersonating a legitimate user, there by infecting the entire cloud thus affecting many customers who are sharing the infected cloud.

C. Proposed System

In the proposed system, the user registers himself with the cloud storage service provider. To enhance security, when the data is uploaded, double-encryption (i.e. by using A.E.S) takes place. Now the data is spitted and transferred across the replica servers. In the servers the encrypted data is converted to binary format and a parity bit is added to it for error detection. In the existing system the trusted third-party auditor is given a copy of the codes to check the data integrity, there arises a situation where the third-party auditor might become a threat to the process. In the proposed system, the data in each server of the multi-server setting is hashed to generate a unique code for each block of data. This code generated for each server is given to the third-party auditor. At regular intervals the data in each server of the multi-server setting is hashed and the third-party auditor checks the integrity by comparing the codes. If any deviation in the code is monitored, then the data might be corrupted or changed by a malicious hacker. The trusted third-party auditor informs that the data is corrupted to the main server and the main server retrieves the data from the server, if the data in the server is also corrupted then erasure coding mechanism is invoked on the servers to regenerate the corrupted data.

► *Advanced Encryption Standard (AES)*

The Advanced Encryption Standard, in the following referenced as AES, is the winner of the contest, held in 1997 by the US Government, after the Data Encryption Standard was found too weak because of its small key size and the technological advancements in processor power. It is a Block cipher, which means that it works on fixed-length group of bits, which are called blocks. It takes an input block of a certain size, usually 128, and produces a corresponding output block of the same size. The transformation requires a second input, which is the secret key. It is important to know that the secret key can be of any size (depending on the cipher used) and that AES uses three different key sizes: 128, 192 and 256 bits. The algorithm is designed to use keys of length 128, 192 or 256. It works on one block of 128 bits at a time, producing 128 bits of cipher text. There are 10 rounds, after an initial XOR'ing (bitwise addition mod 2) with the original key (assuming a key length of 128). These rounds, except for the last, consist of 4 steps (layers), called ByteSub, ShiftRow, MixColumn and AddRoundKey. In the 10th round the MixColumn step is omitted. The 128 bit input is divided into 16 bytes of 8 bits apiece. These are arranged in a 4 × 4 matrix. The ShiftRow and MixColumn steps operate on this matrix while the ByteSub and AddRoundKey steps just operate on the bytes. The important features of this field are that each of its elements is represented by a single byte (8 bits), and one can add and multiply these bytes to get another byte.

How AES Encryption and Decryption Works

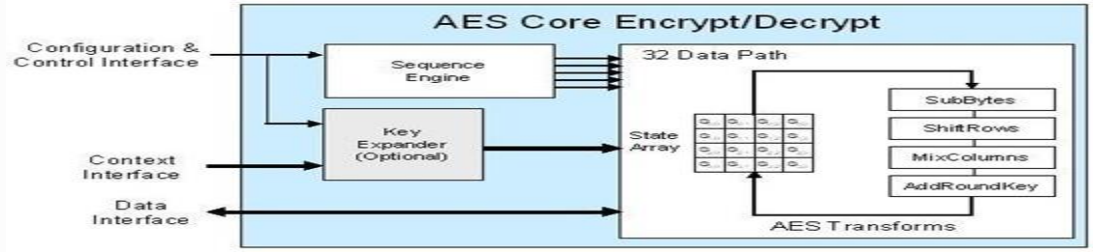

Fig 3 Working of AES

Encryption is the process of converting plaintext to cipher-text (had to understand) by applying mathematical transformations. These transformations are known as encryption algorithms and require an encryption key. AES comprises three block ciphers, AES-128, AES-192 and AES-256. Each cipher encrypts and decrypts data in blocks of 128 bits using cryptographic keys of 128-, 192- and 256-bits, respectively. Symmetric or secret-key ciphers use the same key for encrypting and decrypting, so both the sender and the receiver must know and use the same secret key. All key lengths are deemed sufficient to protect classified information up to the "Secret" level with "Top Secret" information requiring either 192- or 256-bit key lengths. There are 10 rounds for 128-bit keys, 12 rounds for 192-bit keys, and 14 rounds for 256-bit keys. A round consists of several processing steps that include substitution, transposition and mixing of the input plaintext and transform it into the final output of cipher text.

Decryption is the reverse process of getting back the original data from the cipher-text using a decryption key. In Symmetric cryptology, the encryption key and the decryption key could be the same as in symmetric or secret key cryptography, the key can different as in asymmetric or public key cryptography. Some operations which are done in the AES decryption are as follows:

- AES decryption is not identical to encryption since steps done in reverse.
- But can define an equivalent inverse cipher with steps as for encryption.
- But using inverses of each step.
- With a different key schedule.
- Works since result is unchanged.
- Swap byte substitution & shift rows.
- Swap mix columns & add (tweaked) round key.

❖ *Advantages*

- Cost Savings - The user pays for what is used and disengage whenever the user like. There is no invested IT capital to worry about.
- Reliability - With a managed service platform, cloud computing is much more reliable and consistent than in-house IT infrastructure.
- Manageability - Cloud computing provides enhanced and simplified IT management and maintenance capabilities through central administration of resources, vendor managed infrastructure, service level agreement and SLA backed agreements.
- Strategic Edge - Ever-increasing computing resources give the user a competitive edge over competitors, as the time the user requires for IT procurement is virtually nil.
- By striping redundant data across multiple servers, the original files can still be recovered from a subset of servers even if some servers are down or compromised.
- A thin-cloud setting is used where servers only need to support standard read/write functionalities for portability and simplicity.
- Different parameters can be adjusted for the performance-security trade-off.

III. RESULT AND DISCUSSION

Cryptography protects users by providing functionality for the encryption of data and authentication of other users. The graphical representation on discussion of various cryptographic algorithms over security is shown in Fig1.

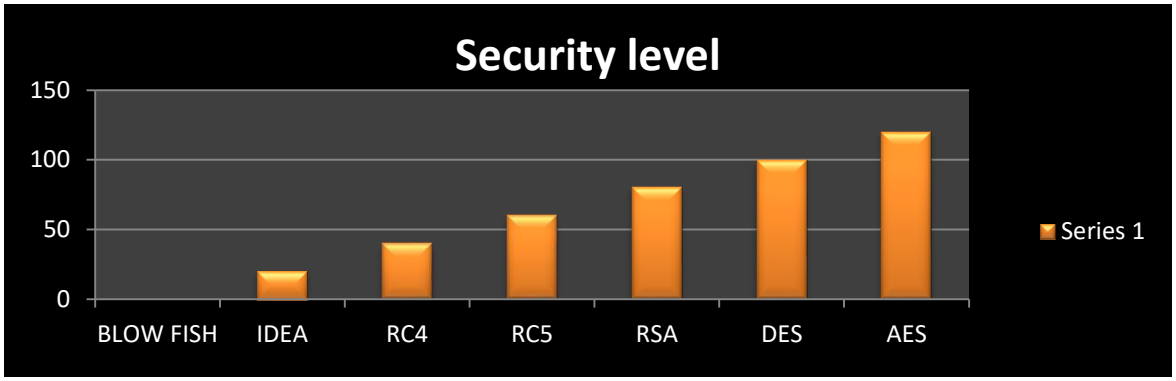

Fig4. Comparison between n cryptographic algorithms

The above fig shows that the AES performs good to secure the data while transferring, Software enhancement may involve providing new functional capabilities, improving user display and modes of interaction, upgrading external documents or the performance characteristics of the system. Implementation is the most crucial stage in achieving a successful system and giving the user's confidence that the new system is workable and effective. Implementation of the modified application to replace an existing one is possible. This type of conversation is relatively easy to handle, provide there are no major changes in the system.

Each program is tested individually at the time of development using the data and has verified that this program linked together in the way specified in the programs specification, the computer system and its environment is tested to the satisfaction of the user. The system that has been developed is accepted and proved to be satisfactory for the user. And so the system is going to be implemented very soon. A simple operating procedure is included so that the user can understand the different functions clearly and quickly. Initially as a first step the executable form of the application is to be created and loaded in the common server machine which is accessible to the entire user and the server is to be connected to a network. The final stage is to document the entire system which provides components and the operating procedures of the system.

IV. CONCLUSION

Cryptography protects users by providing functionality for the encryption of data and authentication of other users. The advanced Encryption Standard (AES) is one of the most popular encryption algorithm used until recently. Today, cloud computing is being defined and talked about across the ICT industry under different contexts and with different definitions attached to it. The core point is that cloud computing means having a server firm that can host the services for users connected to it by the network. Technology has moved in this direction because of the advancement in computing, communication and networking technologies. Fast and reliable connectivity is a must for the existence of cloud computing. Cloud computing is clearly one of the most enticing technology areas of the current times due, at least in part to its cost-efficiency and flexibility. However, despite the surge in activity and interest, there are significant, persistent concerns about cloud computing that are impeding the momentum and will eventually compromise the vision of cloud computing as a new IT procurement model. Despite the trumpeted business and technical advantages of cloud computing, many potential cloud users have yet to join the cloud, and those major corporations that are cloud users are for the most part putting only their less sensitive data in the cloud. Lack of control is transparency in the cloud implementation – somewhat contrary to the original promise of cloud computing in which cloud implementation is not relevant. Transparency is needed for regulatory

reasons and to ease concern over the potential for data breaches. Because of today's perceived lack of control, larger companies are testing the waters with smaller projects and less sensitive data. In short, the potential of the cloud is not yet being realized. When thinking about solutions to cloud computing's adoption problem, it is important to realize that many of the issues are essentially old problems in a new setting, although they may be more acute. For example, corporate partnerships and offshore outsourcing involve similar trust and regulatory issues. Similarly, open source software enables IT department to quickly build and deploy applications, but at the cost of control and governance. Similarly, virtual machine attacks and web service vulnerabilities existed long before cloud computing became fashionable. Indeed, this very overlap is reason for optimism; many of these cloud computing roadblocks have long been studied and the foundations for solutions exist. For the enhancement of technology, and hence healthy growth of global economy, it is extremely important to iron out any issues that can cause road-blocks in this new paradigm of computing.

REFERENCES

1. Abu-Libdeh. H, Princehouse. L, and Weatherspoon. H, RACS: "*A Case for Cloud Storage Diversity*", (2010).
2. Ahlswede. R, Cai. N,. Li. S.Y. R, and Yeung. R.W., "*Network Information Flow*", (2000).
3. Alenis Leon, "*System Analysis and Design*", (1987).
4. Armbrust. M, Fox. A, Griffith. R, Joseph. A.D, Katz. R, Konwinski. A, Lee. G, Patterson.D,Rabkin. A, Stoica. I, and Zaharia. M, "*A view of cloud computing*", (2010).
5. Chris Goode, John Kauffman "*Beginning Asp.Net 1.0 With Visual Basic.Net*" -Wrox Programmer To Programmer.
6. Douglas O.Reilly , "*Designing Microsoft Asp.Net Applications*"-Tata Mcgraw Hill Edition.
7. Goldreich. O, "*Foundations of cryptography: Basic applications*" - volume2, (2004).
8. Goldreich. O, "*Foundations of cryptography: Basic tools*" - volume 1, (2001).

ONLINE REFERENCES

1. http://research.microsoft.com/en-us/people/klauter/cryptostoragerlcps.pdf
2. http://www.appcore.com/types-cloud-computing-private-public-hybrid-clouds/
3. http://www.thoughtsoncloud.com/2014/02/cloud-computing-basics/
4. http://www.tutorialspoint.com/cloud_computing/cloud_computing_overview.htm