

Implementation of Enhanced AES for Secure and Efficient Data Storage in Cloud Environment.

¹N. Velmurugan, ²S. Godfrey Winstler

¹Research Scholar, Saveetha University and AP (SG) / IT, ²Professor / CSE

^{1,2}Saveetha Engineering College, Chennai, India.

velmurugann@yahoo.com, godfrey@live.in

Abstract- Cloud computing is an Internet based platform for computing. Cloud computing is popular for its centralized data storage. This revolutionary computing has enormous advantages on one hand and also has challenges on the other hand. Data Security in cloud is the very important issue that meets immediate attention. This paper proposes a reliable system for efficient data storage by employing tag key based data deduplication and enhanced AES data encryption. This system uses enhanced version of the conventional AES algorithm in order to accelerate the encryption and decryption process. The data deduplication saves storage space by eliminating the redundant data, which improves the efficient usage of the cloud. This proposed system is better than the current cloud security systems as it addresses and satisfies the key disadvantages of the existing systems.

Keywords: cloud computing, Redundancy, Enhanced AES, hash key, data deduplication and cloud security.

I. INTRODUCTION

Cloud computing is based on the concept of internet in which large groups of remote servers are networked to allow the centralized data storage, and online access to computer services or resources. Cloud service offers typically three categories: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Cloud computing has great possibilities and challenges. The first and very important challenge to be addressed is the data security. Cloud computing is using the public networks to store and transmit data which makes the data in the cloud exposed to public and vulnerable. This nature of the cloud computing, leads to security and privacy concerns. Session hijacking and data segregation are the potential and unavoidable security breaches for the cloud users. Cloud is associated with two kinds of resources namely physical and virtual resources. Both virtual and physical resources pose multiple security issues at different levels.

This paper proposes a novel approach for developing a secure cloud access system for the storing and sharing confidential data.

II. RELATED WORKS & CONTRIBUTION

A new cryptographic method is formalized in which the key that is used for both encryption and decryption is derived from the data itself. This cryptographic method is called Message-Locked Encryption (MLE) [1]. Another new deduplication system called RevDedup, is a deduplication method which follows the concept of reverse deduplication. This method eliminates the duplicate data from the new data rather than removing data from the existing data as in the conventional deduplication methods [2]. The cloud security issues such as virtualization security and data security are the key issues restricting the application of cloud computing. These security issues have become an obstacle restricting cloud computing industrial applications [3]. There are few practical methods for securing cloud computing. These practical methods are as follows: private cloud that are configured with enterprise perimeters, public cloud that possess service gateways, provision for content encryption, cloud access broker, session containers, and security virtualization in runtime [4]. AES is the popular and standard data encryption method. In 2001, American National Institute of Standards and Technology published a novel encryption standard called AES (Advanced Encryption Standard). AES is based on a symmetric-key algorithm. In symmetric key algorithms same key is used for both encryption and decryption of the data. The block size is often 128 bits and the key size can be of different sizes such as 128, 192, or 256 bits. AES operates on a state, which is 4×4 matrix of bytes [5].

There is another cloud security system, which is based on the two-way password concept. In this system once the user gets authentication two passwords will be generated. One password is given to the user and the other is given to the cloud provider. The user can only access the cloud if and only both the passwords were given as input. The main disadvantage of this method is that the user has to

trust the cloud provider for no reason and this system is complex [6].

In a recently published paper, a novel cloud security system in which, the data storage and retrieval is performed using the privilege key. The data is first send to the private cloud for storage through privilege key. The data is next send to the hybrid cloud for storage after encryption. The data is encrypted using convergent encryption technique. One important issue with the service platforms for cloud storage is managing the ever-increasing huge volume of data. Deduplication is the promising technique in order to make the data management in cloud computing into a scalable task. This method has captivated more attention recently. Data deduplication is a concept of data compression, which is made possible by finding and eliminating the redundant copies of the data from the cloud storage. This technique is used to enhance the efficiency of storage utilization and can also applied to data transfers through network to reduce the number of bytes that are transferred. Instead of keeping duplicate data copies, deduplication removes redundant data and store only one physical copy and referring other redundant data by an alert message. Cloud backup service is an affordable choice for data protection for personal computing devices. The cloud computing offers centralized cloud management, which created efficiency and cost effectiveness. This offers simple backup storage for disaster recovery, which is always a major consideration for data backup. The IT resources in the cloud have so much redundancy in backup dataset. Hence there is lot of scope for improving the efficiency of its cloud storage [7].

In the existing systems data deduplication is performed through hybrid cloud architecture. The hybrid cloud architecture does not eliminate the redundant data completely. The convergent encryption technique used for encrypting the data is inefficient. The popular AES algorithm can encrypt the block size of 128 bits, which is time consuming. The user uses same privilege key for the storage and retrieval of data. The privilege key is easily predictable by the hackers or intruders.

III. PROPOSED SYSTEM

The proposed system has two tasks namely storage of data in cloud and data retrieval from cloud. For storing the data in the cloud the user first log in using the username and password. The system validates the identification of the user and allows the user to access the cloud. The data to be uploaded is submitted to the system. The system generates a tag key, which is generated based on the content of the uploading file, thus this tag key serves as the fingerprint of the uploading file, which is unique to the file. The tag key generation is followed by deduplication for the removal of redundant data files. The deduplication is performed by comparing the unique tag key of the uploading file with the tag keys of the files that are

already stored in the cloud. If the tag key of the uploading file matches with the tag key of the stored file, an alert message is displayed so that the user can avoid uploading the redundant file. On the other hand, if the hash key of the uploading file doesn't matches with the hash keys of the stored files, the system proceeds to the encryption step. Before encryption a 20 byte unique hash key is generated using the Secure Hash Algorithm (SHA). After the generation of the hash key, the hash key and the uploading file is encrypted using the enhanced AES algorithm. After the encryption procedure the file will be successfully uploaded and stored in the cloud. For the data retrieval the user needs to gain access to the cloud by validating the user's username and password, which is same as that of the storing the data as mentioned. Once the system validates the user, the user can access the cloud and the system will display the files that are stored in the cloud. The user has to request file to the system. The system will process the request after authenticating the user. After processing the file selection the user can get the decrypted hash key and the corresponding decrypted file.

IV. ADVANTAGES OF THE PROPOSED SYSTEM

- The data deduplication is performed based on the data specific tag keys.
- The technique used for deduplication is more effective and simple.
- The data is encrypted using enhanced AES algorithm, which is faster than conventional AES encryption technique.
- The hash keys are used to decrypt the file, thus enhancing the data security.

V. SYSTEM ARCHITECTURE

System architecture provides description of the proposed system, which is organized in a way to support logical reasoning of proposed system's the structural properties. System architecture enlist the building blocks of the proposed system that will work in unison for the successful implementation of the overall proposed system.

This study depicts a system architecture shown in figure1. The components of the proposed architecture are:

- Cloud
- Data load
- Tag key generation
- Data deduplication
- Hash key generation
- Data encryption

- Data retrieval
- Data decryption

Cloud:

The cloud computing platform that is available for the general public without any firewall is called the public cloud. It is cheaper than the private cloud. Our system utilizes the public cloud for the data storage, as it is cheaper. The uploaded files are stored in the database located in the public cloud after the encryption.

Data load:

If the user wants to upload the data to the database the files are first subjected to the unique and file content based tag key generation, which is followed by data deduplication. The deduplication algorithm checks for the redundancy of the files in the database by comparing the tag keys.

Tag key generation:

Tag key generation is the important aspect of our system. This key is based on the content of the file thus making it unique to the file. Tag key of the uploading file is used as search query for the detection of redundant data by comparing with the tag keys of the files that are already stored in cloud.

Data deduplication:

The data deduplication is performed by comparing the tag key of the file to be uploaded with the tag keys of the files that are already stored in the cloud. If there is matching of the tag keys, the file is termed as redundant file. The files can be uploaded and stored in the cloud if and only the tag key of the file to be uploaded does not matches with the tag keys of the files already stored in the cloud. This process is very important as it ensures the efficient use of the cloud space.

Hash key generation:

Hash keys are generated using the SHA algorithm. The hash key is encrypted along with the corresponding file and stored in the cloud. This hash is needed for the retrieval of the file.

Data encryption:

Encryption is the process of converting the normal text into the cipher text. Our system employs Enhanced AES algorithm as the encryption protocol. The uploaded files are first undergoes the process of encryption before stored in the database. The process of encryption serves as the last line of defense against the data piracy. The encrypted files cannot be decrypted without the hash key, which adds further data security.

Data retrieval:

If a registered user wants to retrieve a file from the

database in the public cloud, the user has to get

authorized by the system, by which the user gets the decrypted hash key for decrypting the file.

Data decryption:

The process of converting the cipher text into the readable plain text is called data decryption. In our system we employs Enhanced AES algorithm for the data decryption. Data decryption is the last step for data retrieval.

VI. MODULES

- Service Requester
- Implementation of tag key
- Deduplication
- Enhanced AES Encryption
- Enhanced AES Decryption

MODULE DESCRIPTION:

Service Requester:

The service requester module enables the user to access the cloud. In order to access the cloud the user has to log in to the cloud using authenticated username and password. This module verifies the username and password and allows the user to access the cloud if and only the verification is successful.

Implementation of Tag Key:

The Hash keys are generated based on the content of the data. Hence these keys are unique to the data files. The tag keys are generated once the data file is ready for upload. These keys are used in data deduplication.

Deduplication:

In the deduplication procedure the tag keys of the file to be uploaded and files in the cloud were compared. If the system finds a match, then the system will alert the user regarding data redundancy or else the user is permitted to store the file in cloud.

Enhanced AES Encryption:

Preserving required data has been converting into cipher text for the protection. Enhanced Advanced Encryption Standard is based on a design principle known as a substitution permutation network. Enhanced Advance Encryption Standard is commonly use block ciphers of size 256 bits. By design enhanced AES is faster than the conventional AES as the block size is two times higher. This process has a main goal

is to converting plain text into cipher text from the given array of input information.

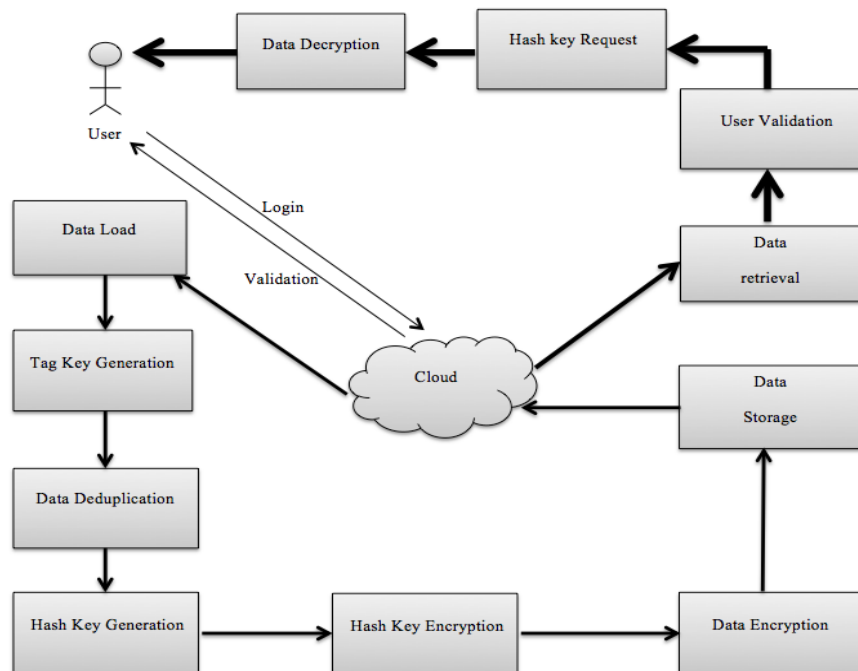


Figure 1- Architecture diagram for the proposed system. Thick arrow represents the data retrieval pathway, normal arrow represents the data storage pathway and the thin arrows represent the login pathway, which is common for both data retrieval and data storage.

Enhanced AES Decryption:

The decryption is the process of converting the cipher text to plaintext. With the help of generated key, cipher text has been changing over the time limit into original information. Unlike enhanced AES encryption, in Enhanced AES decryption steps done in inverse order.

VII. ENHANCED ADVANCED ENCRYPTION STANDARD (AES) ALGORITHM:

Enhanced Advanced Encryption System (EAES) algorithm is characterized as symmetric block cipher that encrypts 256 bits of the text data blocks into cipher keys of different bit lengths namely 128, 192 or 256 bits. This encryption method utilizes the Rijndael algorithm, which was developed by Belgium scientists namely Vincent Rijmen and Joan Daemen. The main advantages of enhanced AES algorithm are it is very simple design, high speed in processing without compromising the memory. Another very important advantage is that the both encryption and decryption are performed using the same hash key.

AES encryption uses the round function. The round function includes of four transformations as follows:

- Byte-sub operation is the non-linear substitution of each byte using the defined substitution table.
- Shift-row operation shifts the rows of the state in a circular fashion.
- Mix-column operation uses the multiplication with fixed polynomial modulo in order to mix the bytes.
- Add round key operation adds round key to the state [5].

Secured hash algorithm (SHA):

Secured Hash Algorithm (SHA) is a cryptographic hash algorithm designed by National Security Agency (NSA) and published by National Institute of

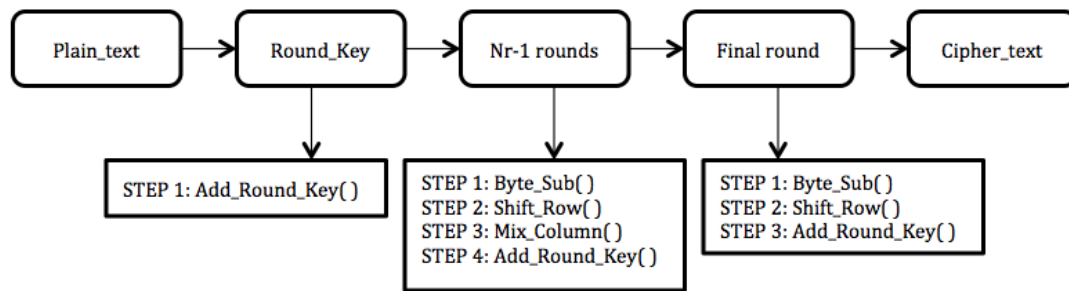


Figure 2-Enhanced Advanced Encryption Standard Algorithm – Encryption.

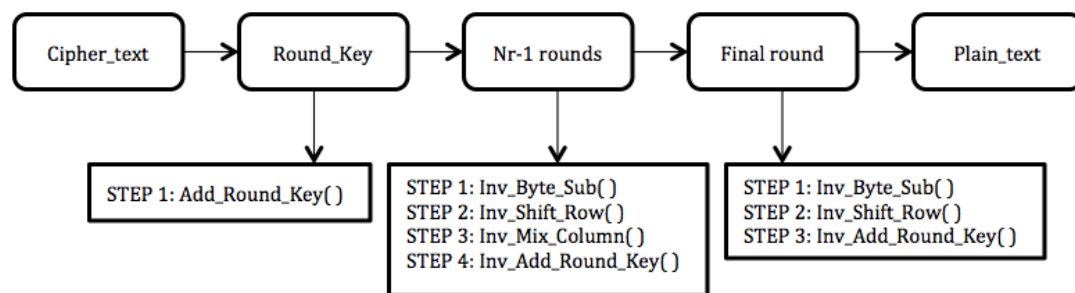


Figure 3-Enhanced Advanced Encryption Standard Algorithm – Decryption.

Standards and Technology (NIST). This algorithm could compress the given information and gives out a hash key of size 160 bits (20 bytes).

VIII. COMPARISON GRAPH

This comparison graph depicts the encryption speed of the AES encryption and Enhanced AES encryption. In the enhanced AES algorithm, as the size of each block is higher than the AES algorithm the total number of blocks (referred as block size in figure 4) is lower in the enhanced AES algorithm. Hence the number of blocks to encrypted is significantly reduced. Thus increases the encryption rate. With the increase in the encryption rate, time taken for encryption is quicker in Enhanced AES encryption procedure when compared to the AES encryption procedure.

IX. CONCLUSION

This proposed system uses tag key based data deduplication technique, which avoids data redundancy saving space ultimately increasing the efficiency of the data storage. This system also uses the standard Enhanced AES algorithm for data encryption thus enhancing overall data security. Finally this also uses the hash key method in data encryption using SHA, which is generated

dynamically in a unique way, which makes the system secure.

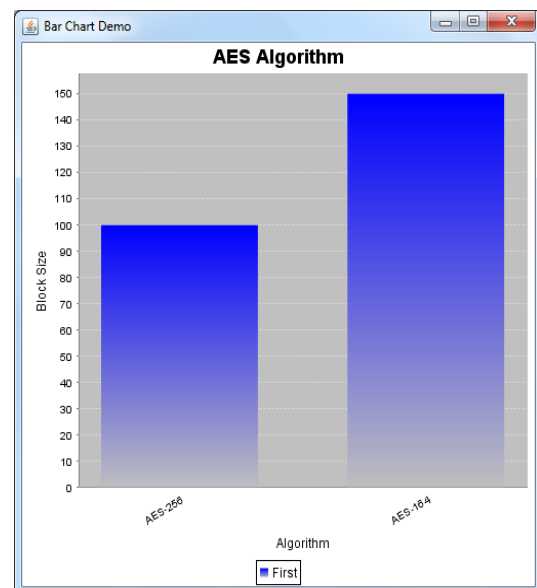


Figure 4- comparison graph between AES algorithm and Enhanced AES algorithm.

X. FUTURE ENHANCEMENT

The problem with the secret keys is when the number of files stored in the cloud increases the number of secret key also increases. Thus the user has to remember many secret keys, which is cumbersome. The concept of key-aggregate cryptosystem is introduced recently. In this method many secret keys could be combined to form a single secret key. While sharing the data, the user can generate a secret key for the sharing file. So that only the file that need to share can be accessed and the rest of the files remain secured [9]. In our future enhancement we like to incorporate modified version this concept in our system. Our modification will include implementing login based secure aggregate key and individual key generation portal for the creation of the aggregate keys and individual keys from the aggregate key. Our current encryption algorithm works fine with the text file, in future we intend to develop and implement an algorithm that could encrypt and decrypt other data formats like images, tables, graphs etc.

REFERENCES

- [1] Bellare.M, Keelveedhi.S, Ristenpart.T (2013), 'Message-locked encryption and secure deduplication', In *EUROCRYPT*, pp 296– 312 .
- [2] Ng, C., & Lee, P. P. C. (2013). RevDedup : A ReverseDeduplication Storage System Optimized for Reads to Latest Backups. *APSys*. doi:10.1145/2500727.2500731
- [3] Luo, X., Yang, L., Hao, D., Liu, F., & Wang, D. (2014). On Data and Virtualization Security Risks and Solutions of Cloud Computing. *Journal of Networks*, 9(3), 571–581. doi:10.4304/jnw.9.3.571-581
- [4] Amoroso, E. G. (2014). Practical methods for securing the cloud. *IEEE Cloud Computing*, 1(1), 28–38. doi:10.1109/MCC.2014.1
- [5] Ankita N., & Sheeja S. (2014a). Design and Implementation of AES Algorithm Using, 2(1), 419–423.
- [6] Brar, Y., Krishan, S., Mehta, A., & Talwar, V. (2014). An Advanced Security - A Two-Way Password Technique for Cloud Services, 3(4), 7–15.
- [7] Li, J., Li, Y. K., Chen, X., Lee, P., & Lou, W. (2014). A Hybrid Cloud Approach for Secure Authorized Deduplication. *IEEE Transactions on Parallel and Distributed Systems*, 1–1. doi:10.1109/TPDS.2014.2318320.
- [8] Zheng, Y., Pieprzyk, J., & Seberry, J. (1992). HAVAL - A One-Way Hashing Algorithm with Variable Length of Output. *Advances in Cryptology - AUSCRYPT '92*, 718, 83–104. doi:10.1007/3-540-57220-1_54
- [9] Chu, C., Chow, S. S. M., Tzeng, W., Zhou, J., Deng, R. H., & Member, S. (n.d.). Supplementary Material for Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage, 1–4.



Received UG degree in physics and PG degree in Computer Applications from Bharadhidasan University, Trichirappalli, India. Completed Master of Engineering degree in CSE at Sathyabama University, Chennai, India. Worked as a project trainee in Flight Dynamic Division at ISRO, Bangalore for IRS-P4 satellite. Working as a Assistant Professor (senior grade) in the department of Information Technology at Saveetha Engineering College, Chennai, India. Has 18 + years of teaching experience in engineering colleges. Area of interest includes Design of Algorithms, Data Structures, Cryptography, Object, Software Design & Implementations. Specialization in Cloud and Bigdata. Doing research on Data Security in Cloud at Saveetha University, Chennai, India. Published 4 international journals. Received the BEST TEACHER award from Saveetha Engineering College. Worked as a Zonal Officer, Zone III, Anna University of Technology, Chennai for all type of examinations related administrative activities to 26 engineering colleges. Life Member of ISTE. Member of UACEE, IAENG and CSTA.



Received Bachelor degree in Computer Science and Engineering from the University of Madras and Master's degree in Computer Science and Engineering from Sathyabama University. Completed the Doctor of Philosophy in Web Mining from Anna University. Research interest includes Web Mining, Semantic Web, Cloud Computing, Data Analytics and Social Networking. Published various papers in international journals and conferences. Has 14 years of teaching experience in various reputed engineering colleges in Chennai. Coordinated and attended various Faculty developments programmes and International conferences. Certified EMC Academic Associate in Cloud Infrastructure and Services. Active member in professional bodies like IEEE, CSI, ISTE, CSTA and IACSIT. Reviewer in various International reputed journals JWS, JDCTA and JCIT. Recognized supervisor in Anna University to guide MS(By research) and Ph.D scholars. Certified EMC Academic Associate in Cloud Infrastructure & Services and Data Science & Bigdata Analytics. Published Compiler Design book in association with YesDee publications.

