Encrypting Data on Cloud Using Advanced Encryption Algorithm

Aditya Singh Kushwah
Assistance System Engineer, Tata Consultancy Services
Mumbai, India
speak2aditya@gmail.com

Abstract - In recent years, cloud computing has been adopted by large sphere of public and private organizations due to its streamline processes, reduced capital, improved flexibility and accessibility, and resource sharing ability. More and more IT industries are taking a voyage to cloud from dedicated data centres. But with the advent of cloud technology comes avant-garde and byzantine issues, primarily, data security issues. While benefits of using cloud are riveting, data security for stow data is still a major topic of apprehension for every IT organization. This paper presents a method to safeguard the data which is stored on the cloud by implementing Advanced Encryption Standard (ASE) algorithm.

Index Terms – Cloud computing; Resource sharing; Data security; ASE algorithm.

I. INTRODUCTION

Cloud computing is defined as a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. Cloud computing is comparable to grid computing, a type of computing where unused processing cycles of all computers in a network are harnesses to solve problems too intensive for any stand-alone machine. In cloud computing, the word cloud (also phrased as “the cloud”) is used as a metaphor for “the internet,” so the phrase cloud computing means “a type of Internet-based computing,” where different services – such as servers, storage and applications – are delivered to an organization’s computers and devices through the Internet.

Data stored on cloud is far more accessible, cost effective, nimble, and amiable to environment than the conventional way of storing the data as long form of manuscripts. But even if a small bit of data which is stored digitally is left exposed can create turbulence as it is at potentially high risk of being pilfer. Cloud is one such platform where user/organisation can store their data on the space provided by the vendors of the cloud. Cloud helps organisations to proliferate their business in diverse fronts and not worrying about the infrastructure prospects. Also as the business grows, the ever-changing demand for the business increases and cloud support high flexibility in choosing the resources as and when needed according to the need to the consumers. The cloud is based on “pay as you go” model.

Even though cloud possesses many advantages, but the biggest threat persist to the data that it carries as it is susceptible to attack if proper security measures are not taken. As cloud holds a large amount of data, the vendor of the cloud should equally be cautious to the data which may be venerable to the attack. Stringent algorithm and proper encryption techniques are some steps that can be taken to protect the stored data. Encipherment is one of the security mechanisms to protect information from public access [1]. Encryption hides the original content of a message so as to make it unreadable for anyone, except the person who has the special knowledge to read it [2].

II. CLOUD SERVICE MODELS

Cloud computing vendors offer their services according to three fundamental models. They are: Infrastructure-as-a-Service (IaaS): Infrastructure-as-a-Service is the first layer and foundation of cloud computing. Using this service model, you manage your applications, data, operating system, middleware and runtime. The service provider manages your virtualization, servers, networking and storage. This allows you to avoid expenditure on hardware and human capital; reduce your ROI risk; and streamline and automate scaling. According to a 2011 article released by Venture Beat, Some of the biggest names in IaaS include Amazon, Microsoft, VMWare, Rackspace and Red Hat.

Platform-as-a-Service (PaaS): Platform-as-a-Service is the second layer in the service model. Cloud service providers offers the prerequisites like Operating Systems, programming language environment, web servers, database etc. and the application developer can develop and test the application. The extra cost of buying and maintaining the software can be considerably cut
down as cloud providers manages these services. Microsoft Azure and Google App Engine are examples of some popular Platform-as-a-Service.

**Software-as-a-Service (SaaS):** Software-as-a-Service is the last layer in the service model. In the business model using software-as-a-service (SaaS), users are provided access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes referred to as "on-demand software" and is usually priced on a pay-per-use basis or using a subscription fee. The price depends on the use of resources by the users. It can be scaled according to the need of the user.

**III. DEPLOYMENT MODELS**

In order to cut down the capital expenditure and operating cost, many organisations are moving to cloud. However, with Cloud providing many positive prospects for business, it can also be expensive as cloud can bring high security risk. So, before moving to cloud, the organisation should be clear with the business needs as to which model will be the felicitous for them. National Institute of Standards and Technology (NIST) recommends four cloud deployment models.

a) **Private Cloud:** Private clouds are specifically based for a particular company/organization. Installing a private cloud is not much cost-effective as it has to be bought, built, and managed by the organisation. But it provides a high security aspect to the data stored. As private clouds are only dedicated to the specific user, the security challenges faced by the cloud are taken care in this model. As organisation still have to buy, maintain, upgrade the cloud resources on a regular basis, the private cloud have come under criticism for lack of economic model for which cloud technology is known.

b) **Public Cloud:** As compared to private cloud, public cloud model is cost-effective. In public cloud deployment model, services and infrastructure are provided by the cloud vendors to the user as per the need of the user. It is a flexible model and the resources are allocated according to the demand proposed by the user according to its need. Public cloud works on pay-as-you-go model. Unlike private cloud, public cloud can have security issues as its services are based on network that is open for public use. The infrastructure and security are provided by the vendors and once the data is stored on the public cloud, the user is not aware about the possible measures that are taken to safeguard the data from unauthorised access. This model helps to cut capital expenditure and bring down operational IT cost. For example, Google Cloud Platform is a set of modular cloud-based services that allows user to create anything from simple websites to complex applications.

c) **Hybrid Cloud:** As name insinuates, hybrid means from varied source. Hybrid cloud have configuration of two or more cloud (private, public or community) and offers the facility of different models in a single package. Hybrid cloud can offer its user features like scalability, cost efficiency, security, and pliability.

A hybrid cloud service crosses isolation and provider boundaries so that it can't be simply put in one category of private, public, or community cloud service. It allows one to extend either the capacity or the capability of a cloud service, by aggregation, integration or customization with another cloud service. For example, an organisation can rely on private cloud model for enhanced security, and for large capacity need organisation can use public cloud computing model. Hybrid cloud models can be implemented in a number of ways:

1) Separate cloud providers team up to provide both private and public services as an integrated service
2) Individual cloud providers offer a complete hybrid package
3) Organisations who manage their private clouds themselves sign up to a public cloud service which they then integrate into their infrastructure.

d) **Community Cloud:** A Hybrid Cloud is any combination of Clouds. The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, or compliance considerations). It may be managed by organisations or by a third-party and may be located on premise or off premise. Community clouds are a hybrid form of private clouds built and operated specifically for a targeted group. These communities have similar cloud requirements and their ultimate goal is to work together to achieve their business objectives.
For example, an enterprise can deploy an on-premises private cloud to host sensitive or critical workloads, but use a third-party public cloud provider, such as Google Compute Engine, to host less-critical resources, such as test and development workloads. To hold customer-facing archival and backup data, a hybrid cloud could also use Amazon Simple Storage Service (Amazon S3). A software layer, such as Eucalyptus, can facilitate private cloud connections to public clouds, such as Amazon Web Services (AWS).

IV. CHARACTERISTICS OF CLOUD COMPUTING

NIST (National Institute of Standard and Technology) recognizes the five essential characteristic of cloud computing as follows [3]:

a) **On-demand self-service**: A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

b) **Broad network access**: Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

c) **Resource pooling**: The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

d) **Rapid elasticity**: Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.

e) **Measured service**: Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

V. BENEFITS OF CLOUD COMPUTING

a) **Flexibility**: The main reason that many organisations are moving to cloud is due to its flexible nature of providing the resource according to the need of growing business. For example, if an organisation needs more bandwidth then the current use, the company can anytime demand for the increase bandwidth as per their needs.

b) **Recovery from disaster**: The recovery of data when the data is on cloud is faster as compared to the data stored on other the physical devices in case of any disaster. The data can be recovered at a rapid rate as compared to the one who are not using cloud technology.

c) **Resource sharing**: In cloud, resources are shared among multiple users. One way of achieving this sharing is through virtualization of resources. Virtualization is a way to create a virtual image of a resource, like server, storage device, network or even an operating system. Virtualization provides the vendor the facility to share same resource among multiple consumers.

d) **Ubiquitous availability**: As cloud technology need internet to operate, it gives the consumer a flexibility to use its functionality from around the globe at any point in time when needed. This enhanced capability of cloud gives an edge to the cloud when compared to its counterpart.

e) **Amiable to environment**: As multiple users are bestowed with same resource, which is possible through virtualization, it limits the individual consumption of same resources by different users which in turn effectively cut the amount of detrimental affect considerably.

f) **Accessibility**: When the data is stored on the network, the data is readily available to the user. Even in case of the machine being damaged or stolen, the data can still be accessed through cloud, if stored, which gives the reliability to the users to use cloud.

g) **Cost effective**: Cloud computing is one of the most cost-effective solution in comparison to a standalone system. Purchasing software for standalone systems can be costly due to its high maintenance, future up-gradation, licensing fees, of same software for multiple systems. Cloud technology greatly reduces these expenses. Further, there are facilities that cloud offer, namely one-time-payment, pay-as-you-go, which enhances the robustness of cloud.

h) **Storage**: Cloud provides you almost unlimited space for storage of data. The space can be expanded as per the need of the user.
i) **Backup and Recovery:** Since the data is stored on the cloud, the backing and recovering up of data is rationalized. At times the cloud is itself used as a backup storage place holder for sensitive data. Cloud providers itself provides an efficient solutions for backup and recovery of data.

VI. SECURITY CHALLENGES

Major concerns and challenges for user while storing the data on the cloud are security and privacy. Security is considered as one of the biggest barrier in the road to success for cloud computing. Data breaches and cloud service abuse rank among the greatest cloud security threats, according to Cloud Security Alliance. Though cloud provides very many benefits to the organisation but the security provided by the cloud vendor should be the priority of the user using the services. The security challenges for cloud computing are substantial. One of the challenge of cloud computing is location of the stored data. When the cloud provider moves user data to other servers located in different countries, it might bring serious concern to the security of data and availability as different security norms apply to the data for different countries about which the user is unknown. Security of personal data is one of the concern of cloud computing. Proper security should be implemented to hinder any pilfer of user’s personal data. As the cloud data travels through the network (Internet) which is sometimes unreliable as it is open, the data should be encrypted as per the standard protocols that are developed for the secure transmission of data over the network. There can be many threats while the data is being travelling from one point to other in the network. To name few: IP spoofing, port scanning, man-in-the-middle attack, phishing, eavesdropping, packet sniffing, Distributed Denial of Service attack (DDoS) and many others. IP spoofing, also known as IP address forgery or a host file hijack, is a hijacking technique in which a cracker masquerades as a trusted host to conceal his identity, spoof a Web Site, hijack browsers, or fain access to a network. A Packet Sniffer is a program that can record all network packets that travel past a given network interface, on a given computer, on a network. A man-in-the-middle attack is a type of cyber-attack where a malicious actor inserts him/herself into a conversation between two parties, impersonates both parties and gains access to information that the two parties were trying to send to each other. A denial of service (DoS) attack is a malicious attempt to make a server or a network resource unavailable to users, usually by temporarily interrupting or suspending the services of a host connected to the Internet. When an organization stores its data on the cloud, the physical access to the server on which the data is stored is relinquished due to which there is abiding threat of inside as well as outside attack on the confidentiality and integrity of data. As public cloud is open and the space on which the data is stored is shared by many other users who can be the competitor or hackers. The user’s data is at constant risk if proper and secure segregation of space is not made. As many companies stores there data on the cloud servers, hacker can gain access to the stored information via single attack. The process is called hyperjacking. When the client data resides in the cloud, there is a vital security hurdle. Now the question that arises is that how long the data has to remain in the cloud. Till the client object refers the data, the data is available in the server. There is a chance that cloud provider might retain the information, though client is no longer accessing the data. User has to make sure that their data has to be destroyed or no longer visible in cloud provider domain, when he is migrated or terminated the service from the cloud provider. For this purpose data sanitization has to be done. Data sanitization also applies to backup copies made for recovery and restoration of service and also residual data remaining upon termination of service. [4]

If the data is not properly destroyed after the user no longer uses the services, it is possible to recover the data back. Even though cloud provides minimal cost and high-efficiency for data storage, equal and potentially higher threat on data still exist which make user sceptical for storing data on cloud. As the number of users are increasing so the number foes proliferates for attacking the data. The only question comes to the user who uses the facilities provided by cloud vendor is what potential steps are being taken to safeguard the stored data?

VII. ADVANCED ENCRYPTION STANDARD

Advance Encryption Standard (AES) algorithm, also referenced as Rijndael, is a method used for encryption of electronic data. Advanced Encryption Standard algorithm was developed by two Belgian cryptographers- Joan Daemen and Vincent Rijmen. AES is based on the principle known as Substitution Permutation network (SP-network) which means there will be a series of linked mathematical operations in the block cipher algorithm. [5] AES, unlike other public-key cipher, uses same key for both encryption and decryption of data and hence AES uses symmetric key encryption key algorithm. It has a robust algorithm and longer key length which makes ASE more secure and stronger than its predecessors like DES.
and 3DES. AES was introduced to replace the DES [6]. Whenever a user stores its data on the cloud AES algorithm encrypts the data and then stores it. And when the user retrieves the data back, the AES decrypts the encrypted data using the same key which it previously used for encryption and provide the original data to the authorized user.

AES has a fixed block size of 128 bits and can select a key with sizes 128, 192, 256 bits which are in the multiple of 32 bits with a minimum of 128 bits and maximum of 256 bit. Block is a reference to the bytes that are processed by the algorithm.

For conversion of plain text to cipher there are number of rounds through which a plain text goes before it finally gets converted to its respective encoded format. Number of rounds depends on the size of the key length that is being used for encryption. Keys with length 128, 192, 256 bits have 10, 12, 14 rounds respectively.

Structure of Advance Encryption Standard Algorithm:

1) Round keys are extracted from the cipher key.
2) Block data (Plaintext) is loaded into the state array.
3) Add the round keys to the respective state arrays.
4) According to key selected, (128, 192, 256 bits) the number of rounds, excluding one, are performed for manipulating the state.
5) Perform the final round of state manipulation which is different from rest of the rounds.
6) The final state that has been manipulated is the desired encrypted block.

AES algorithm consists of following steps for encryption of data:

Let us take an example of encrypting a block of size 128 bits with a key of size 196 bits. As the key size is 198 bits, the number of rounds will be 12. Although there are twelve rounds, thirteen keys are needed because one extra key is added to the first state array before the rounds start. The descriptions of steps involved in the algorithm are as follows:

a) KeyExpansion – Number of round keys are derived from cipher key using Rijndael’s key schedule. A separate, unique key of 128 bit is required each time a round is performed.

b) InitialRound-
   i) AddRoundKey – The key that is extracted from the cipher key is conglomerated with each byte of state using XOR operation.

c) Rounds-
   Multiple operations are applied on state array and each time a new state array is generated with the execution of round. In the first eleven rounds, the four operations are performed in the order listed below. In the twelfth round, all operations are performed excluding the MixColumns operation. Following are the steps required for computing state array for each round:
   i) SubBytes– In this step, each byte in state matrix is replaced with a SubByte of 8-bit Substitution box or Rijndael S-box. Substitution box or Rijndael S-box is a matrix used in the Rijndael cipher which is used to perform substitution.
   ii) ShiftRows - ShiftRows operates on each row of the given state array. Each row of the state array is rotated to its left by certain bytes. The first row of the state array is left unchanged. The second, third, and the fourth rows are rotated left by one, two, and three bytes respectively. For blocks with size 128 bits and 192 bits, the shifting patterns are same but with 256 bit it is different. With 256-bit block, the first row is unchanged and the shifting for the second, third and fourth row is 1 byte, 3 bytes and 4 bytes respectively.
   iii) MixColumns– It is a mixing operation which operates on each columns of the state, combining the four bytes of each column. MixColumns is used for mixing up of the bytes in each column separately. The input byte in MixColumn is further scramble up. It is the operation that mixes the bytes in each column by the multiplication of the state with a fixed polynomial matrix [7].
   iv) AddRoundKey – In this step, the sub-key derived from the Rijndael’s key schedule is combined with the state. The size of the key is same as the size of the state. In AddRoundKey transformation, a roundkey is added to the State by bitwise Exclusive-OR (XOR) operation
Every time a key is added to a state, a new state is formed (Key + State = New State). AES algorithm consists of following steps for decryption of data:

Decryption is simple after understanding the encryption process. It is basically just the inverse. The algorithm was designed for all the steps to be invertible so decryption is basically like doing everything backwards. Therefore, for decryption starts at the last round and the last round key. When processing each round do the process backwards. So, the round key is added first to the last round. Addition is own inverse, which is nice. Then the MixColumn step is applied. The MixColumn step is applied to all rounds except the last one. Also the inverse MixColumn table is used. This table is generated with another matrix similar to the way the MixColumn table was generated. The difference is that there are no short cuts to generate the table. Therefore, the matrix multiplication needs to be performed in the field GF(2^8). All the shifts are done backwards as well. So, instead of shifting left, we shift right. Lastly, the S-Box is applied using the inverse S-box table. The inverse S-box table can easily be generated by taking the S-box value at some row and column index and assigning the row and column to the inverse S-box value at the inverse S-box index defined by the S-box value. For example, the S-box has a value of 0x00 at index 5,2. This translates to the inverse S-box having the value 0x52 at index 0,0. After all the rounds have been completed in opposite order the final state will contain the original plain text.

VIII. RESULTS

In this section, we will design the algorithm which will work on the stated steps discussed before. Below is the Advanced Encryption Standard algorithm:

State (S) = State is a two-dimensional array of bytes on which the AES algorithm’s operations are performed.

Block Size: 128bit
Key length (K): 128, 192, or 256 bit
Number of rounds (r): 10, 12 or 14.

AddRoundKey(S, K_p) //First Step

for p = 1 to (r - 1) //N steps
SubBytes(S, S-box)
ShiftRows(S)
MixColumns(S)
AddRoundKey(S, K_p)
end for
SubBytes(S, S-box) //Last Step
ShiftRows(S)
AddRoundKey(S, K_r)
//MixColumns function is not included in the last step

Y = S
The resultant Y that is obtained is the required encrypted data.

Encryption

Plain Text: Hello! I am AES encryption algorithm.
Encryption key: Secret
Key Length: 128 bits
Result (in base64):
ffurt+lJpLRwpCd7JbontVgDIKlpoMEULekdPvyi7VUIL7DdqxUSr+iFgcUJlqmS

Decryption

Encrypted Text: ffurt+lJpLRwpCd7JbontVgDIKlpoMEULekdPvyi7VUIL7DdqxUSr+iFgcUJlqmS
Decryption Key: Secret
Key Length: 128 bits
Result (Plain Text): Hello! I am AES encryption algorithm.

IX. DISCUSSION AND COMPARATIVE STUDY

AES uses key of different lengths. Every bit generates different possibility of keys.

1) 128 bit: $2^{128} = 3.4 \times 10^{38}$ number of keys (approx.).
2) 192 bit: $2^{192} = 6.2 \times 10^{57}$ number of keys (approx.).
3) 256 bit: $2^{256} = 1.1 \times 10^{77}$ number of keys (approx.).

A PC that tries $2^{55}$ keys per second needs 149.000 billion years to break AES.

There are number of reasons which makes AES more secure than its counterparts. The following table gives a detailed comparison among three different algorithms (AES, DES, and RSA).

<table>
<thead>
<tr>
<th>Factors</th>
<th>AES</th>
<th>DES</th>
<th>RSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developed</td>
<td>2000</td>
<td>1977</td>
<td>1978</td>
</tr>
</tbody>
</table>
Key Size | 128, 192, 256 bits | 56 bits | >1024 bits
---|---|---|---
Block Size | 128 bits | 64 bits | Minimum 512 bits
Ciphering & deciphering key | Same | Same | Different
Scalability | Not Scalable | It is scalable algorithm due to varying the key size and Block size. | Not Scalable
Algorithm | Symmetric Algorithm | Symmetric Algorithm | Asymmetric Algorithm
Encryption | Faster | Moderate | Slower
Decryption | Faster | Moderate | Slower
Power Consumption | Low | Low | High
Security | Exellent Secured | Not Secure Enough | Least Secure
Deposit of keys | Needed | Needed | Needed
Inherent Vulnerability | Brute Forced | Attack Brute Forced, Linear and differentia l cryptanalys is attack | Brute Forced and Oracle attack
Key Used | Same key used for Encrypt and Decrypt | Same key used for Encrypt and Decrypt | Different key used for Encrypt and Decrypt
Rounds | 10/12/14 | 16 | 1
Simulation Speed | Faster | Faster | Faster
Trojan Horse | Not proved | No | No
Hardware & Software Implementation | Faster | Better in hardware than in software | Not Efficient
Ciphering & Deciphering Algorithm | Different | Different | Same

X. CONCLUSION
As cloud is an emerging technology and many organisations rely on cloud for storing its valuable data, the security provided by the vendors should be one step further. A secure environment gives the user a sense of protection for using cloud resources for storing their data. Security of cloud counts on proper and correct implementation of algorithms. Control over the data is relinquished once the data is uploaded on the cloud. To provide a trusted environment to the user, proper data security measures should be implemented on the stored data so that any attempt to intrude the cloud and hamper the authenticity of user’s data can be made foil. Even if the data is leaked accidentally, the decryption of encrypted cloud data will be difficult. Therefore, one way of providing the security is using the implementation of Advanced Encryption Standard algorithm. Our proposed work provides a way to safe guard the stored data on cloud by correctly implementing AES algorithm as this algorithm uses symmetric key to encrypt data and simultaneously decrypt it using the same key. This combination of keys to encrypt and decrypt makes a secure way to store the data on cloud. Further it also avoids any tampering of data and maintains the integrity of the data.
XI. REFERENCE


[7] Chih-Pin Su, Tsung-Fu Lin, Chih-Tsun Huang, and Cheng-Wen Wu, National Tsing Hua University, “A high throughput low cost AES processor” IEEE Communications Magazine 0163-6804/03 © 2003 IEEE.

