An analysis of widely used data security systems in cloud based computing, scope and challenges

Gurpreet Singh[1], Ashoke kumar mahato[2]
[1]*Assistant professor, Department of Computer Science, St.Xavier's College, Ranchi.*
[2]*Ex. H.O.D., Department of Mathematics and Ex. Dean of Science, D.S.P.M.U., Ranchi.*

***Abstract*: Along with the growing convenience and ease of storage of data on clouds and its OTG access, challenges of its security and confidentiality are also increasing with the same pace. Traditionally, the attempts of malicious access of data stored on clouds are considered as the major risk. At present the severity of this risk is becoming more complex. This is because, with the enhancements in technology, the number of cloud service providers is rising and this rising number is causing difficulties to administrate and monitor the overall scenario. When we think about the consequences of this, we understand that the service providers too can access the data in a malicious way and current means of security has no provisions to deal with this situation. At present SaaS (Storage as a service) is most popular utilization of cloud computing and current security provisions are limited up to the identification of user by some credentials like user id and passwords or biometric means in some cases. The valid users may be confirmed with this but the data stored on cloud is fully available to the service providers and they can access it without restrictions. In our study, we are analyzing this unseen aspect of data security and suggesting some ways to safeguard the data from that particular kind of security threat.**

Keywords: SaaS, Biometric security, OTG access.

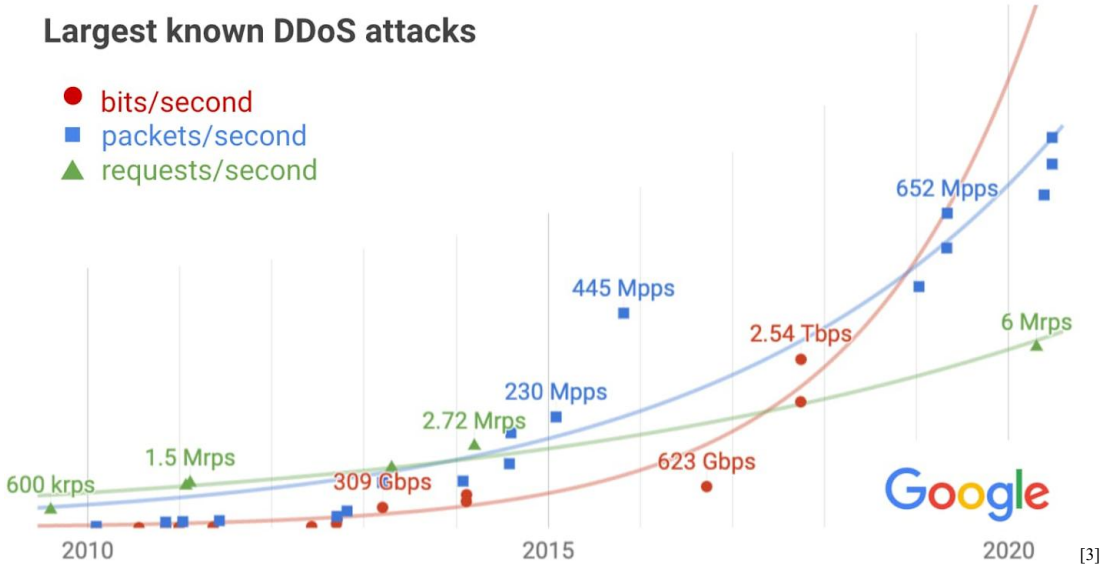
I. INTRODUCTION

The world of computing is rapidly changing by the effects of evolution of a newer aspect of computing that is cloud computing. Its history is traced back to 1950s when an American computer scientist & psychologist Dr. Joseph Carl Robnett Licklider introduced earliest ideas of such type of global network based computing.[1] However, modern cloud infrastructure started to take its shape in early 2000s when world wide web was being spread through the globe. Later in 2002, Amazon web services started cloud based storage and computing services. During the same time, google also introduced the Google Apps Suite (At present Google workspace) which is a collection of SaaS productivity applications. Along with the growth seen in this particular area of computing, the challenges in its path are also seen increasing with the same rate. Today, cloud computing became an integral part of everyday computing. Almost, every internet user is familiar with the use of online available data storage for uploading and storing their files, images and other kinds of data. This type of computing practice is Saas and more precisely StaaS (Storage as a service) which is one of the most important aspects of cloud computing.

Growing utilization of Cloud computing as compared along with Grid and Distributed Computing[2]

II. RESEARCH AND ANALYSIS

When we focus on other side of this convenient and easy to use computing form, we observe a lot of challenging factors present over there. Most important of those is probability of malicious access of data. Till date, the security mechanisms to safeguard the stored contents are limited up to the use of traditional login-ids and passwords. Some of the service providers may offer biometric security which will further add to it. However, all these mechanisms are limited to the authentication of users accessing the service. As far as the security of data is concerned going beyond the conditions of service agreement and either humans or software related to the cloud organization itself are added to the list of potential attackers, all the present mechanisms of security are failed.



[3]

Two most prominent types of attacks related to the case studies in this regard should be pointed first to understand the objective of the our research:

II.i. Malicious service agent:

A malicious service agent is able to intercept and forward the network traffic that flows within a cloud.[4] This service agent can exist in form of a program pretending to be a genuine service program with altered behavior. In some cases, malicious service agent exists as a human who pretends to be a genuine member among the people working with the organization providing cloud based storage service. This type of security threat cannot be handled be conventional login id and password based security methods as the data is always available to organization members. Here, some different kinds of extended security ideas and mechanisms should be discovered, developed and implemented to ensure the safety of data.

II.ii. Malicious insider:

This kind of threat is human driven. The said term is used to point a working member inside the organization of somehow related to the organization. Malicious insiders usually have extended privileges to access the data and they also have access to the premises of the service provider firm. Sometimes malicious insiders are former employees attacking with almost everything known to them. Sometimes, a malicious insider can cause extremely serious damage as they may have privileges of administrative levels[5].

Here, with the above explanation we can observe and understand that as far as the security of data is concerned, conventional methods like user Ids and passwords etc are not sufficient. After seeing above shortcomings and complexities in cloud based storage, we need some mechanism to store data with added measures of security. The data uploaded onto the clouds is always available to the service providers and they can misuse it due to obvious reasons. This kind of threat we discussed under the title “malicious insider”.

III. Study of some prominent technologies:

Our requirement is to encrypt the text data in some complex and specific way to safeguard it and encode its meaning because of its availability to the insiders specified above. Some common ways of encryption of data at present should be outlined first to understand the pros and cons. After that we can propose some solution in order to suggest the lacking features in the current means. Some important approaches of data encryption used for cloud storages are

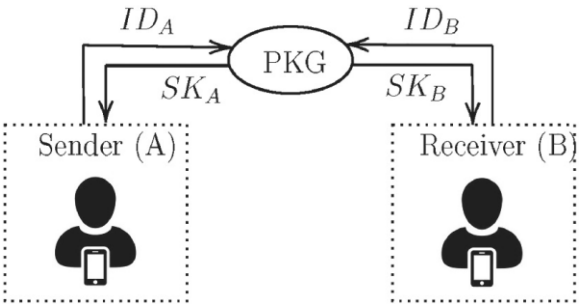
III.i Identity based encryption (IBE):

It was first proposed by an Israeli cryptographer Adi Shamir in 1984.[6] This method is considered as a type

of public key encryption approach in which the public key of a user is some unique information about the identity of the user. Any sender who can access the available portion of the working system is allowed to encrypt a message using identity as a key. For an instance, the message can be encrypted using user's email address as a key. Further the receiver will obtain its decryption key from a central authority. The steps of IBE can be summarized as:
The process of encryption and decryption proceeds as follows:

Step 1. User A gets a message to be sent to user B first. This message is treated as plaintext. While using Identity based encryption, User A uses Id of User B, say idUserB and PKGs public key, say pkPKG to encrypt the plaintext to get cipher text. Then this cipher text is transmitted towards the destination. Here we should understand that pkPKG and idUserB, these two parameters are pre known to UserA. Due to this, in order to prepare the message as well as to encrypt it, User A has no need to establish any communication or to get any prior message from User B to imitate the process.

Step 2. User B gets the encoded cipher text from User A. User B here needs to communicate with the PKG to obtain the private key because however the message is encrypted from the side of User A, the private key from PKG is required to decode it to get the message back in original form. For this UserB is required to send idUserB to PKG in order to prove the authenticity. Once the authenticity is confirmed, PKG will transmit the private key of User2, say skidUser2. This communication is performed via a secured channel and it may be based upon email based system in which some enough level of assurance about the authenticity of UserB is required for confirmation. Sometimes, UserB may be asked to contact PKG in person for highest level of authenticity and to get the skidUserB hand to hand.



Role of PKG between sender and receiver[7]

Step 3. UserB after getting skidUserB can use it to get the message received from UserA in decrypted form with is same as the plaintext prepared by UserA initially.

Here, the problem that we considered for analysis remains. Any central authority which is termed as PKG that generates secret keys for every user is certainly required to be trusted. Any leakage from the authority levels of PKG or any kind of key theft or unauthentic key grants will cause the malicious access to the data.

III.ii Hashing algorithm (SHA-256):

In simplest possible terms, hashing refers to scrambling, mixing
and rearranging data in such a form that it cannot be restored back and reproduced in its original format without knowing some specific system of its encoding. Hashing algorithms are based upon this concept of scrambling and mixing the data. In cloud based storage, the most widely used hashing algorithm is SHA-256 which is commonly known as secured hashing algorithm. This algorithm has many such qualities that make it very suitable for encryption of data to be stored on cloud storages. One instance in this regard is that, SHA-256 always generates a 256 bit hash and this length is independent of the length of input data. It is needed to mention here that hashing is different from cryptographic processing and hashing generates one sided code form which is not retransformed to the original form back again.

Steps of the SHA-256 Algorithm to explain its workability are:

III.ii.i. Getting the input message:

- The input message from the user should be of length multiple of 512 bits. If it is not, then it is padded to make its length same as any multiple of 512 bits.
- After making the message length as multiple of 512 bits, a single "1" is appended to the input and then it is followed by enough "0" bits to make the length of the message 64 bit less than the next or other multiple of 512.
- The original message length (in bits) is appended as a 64-bit binary number.
- After the completion of above 3 steps, the processing results in a message block that is a multiple of 512 bits in length and the input message is considered to be ready.

III.ii.ii. Process of Initializtion:

- In the step of initialization, eight 32-bit registers viz. a, b, c, d, e, f, g, h are initialized with obtained constant values after processing.
- To obtain these constant values, first 32 bits of the fractional parts of the square roots of the first 8 prime numbers are obtained and used as constants.

III.ii.iii. Dividing message into independent portions and processing the parts in round by round manner:

- Because above steps yield the message with length multiple of 512 bits, independent message portions are obtained by dividing the overall length of padded message into 512 bit length parts.
- Every 512 bit portion obtained through above step is processed through 64 rounds of encoding processing in which, at every round the smaller 512 bit portion is divided into 16 number of 32 bit message words(here a word is a binary word of 32 bits).
- All these 16 words are combined together to create a message schedule. This message schedule is process along with specific functions for hash based processing. Bitwise operations, logical functions, and modular addition are applied to the message schedule, round constants, and hash variables.
- At each successive round, these hash variables are updated and new hash value is obtained by combining the result of the current round with the has value from the previous round.

4. Final Hash Value:

- Once the processing is completed, the contents of all 8 bit registers viz. register a,b,c,d,e,f,g,h that were initialized in step 5 are combined together. The finally obtained hash value is therefore the concatenation of all eight 32 bit registers.

Hashing system enhances the security and integrity of data up to a great extent. This can be simply seen in the enhanced security Enhanced Password Security. Hashing protects passwords by transforming them into a one-way hash value, making it extremely difficult to recover the original password even if the database is compromised. Further, it's another advantage is seen in a proven way to ensure the integrity of the crucial data where, the hash value of original data and hash value of retrieved or received data can be compared together in order to ensure that there is no alteration and the integrity is intact. Additionally it is a versatile algorithm that can be used and implemented in a variety of applications like digital signatures, SSL certificates, password managements, block chain technologies and so on. However, despite these merits of the approach, it is sometimes said to be prone to collision. This is because in modern era, since use of technology and digital data is becoming extremely common and due to this the size of databases are becoming very large. In such a situation, it became probable to get some such matching entities in the databases for which the result after hashing may become similar and such a scenario is considered as collision. Another disadvantage of hashing system is that, it when used with large sets and combinations of information, the result will become tedious and time consuming due the required algorithmic processing at different stages.

IV.Conclusion: Despite the availability of several tools, techniques and systems available to safeguard the data stored onto the clouds, the user needs to trust the service providers for the safety of sensitive information belonging to him. This breach of trust is however uncommon in the modern professional era but until there is a loophole there is a chance. Due to this, the need of such an algorithm exists that can encrypt the data in a way that the system of encoding and decoding is known partially by the user and the service provider. The idea is to encode the data upto some extent and the user's site and afterwards it is going to be uploaded where it is again being encoded by the cryptographic and other encoding systems offered by the providers. This 2 level data security

system will help the data remain protected from the malicious insiders and the cases of breach of trust. The upcoming time is will definitely see increasing utilization of cloud storages and increasing amount of data flow towards the online storages. The security system partially shared among the user and the provider is necessitate of the coming era.

REFERENCES

[1] https://www.ibm.com/think/topics/cloud-computing#:~:text=The%20origins%20of%20cloud%20computing,discussing%20an%20Intergalactic%20Computer%20Network.
[2] https://www.researchgate.net/figure/Popularity-of-Cloud-Computing-as-reported-by-Google-trends_fig1_267719048
[3] https://www.spiceworks.com/it-security/network-security/news/how-google-foiled-the-largest-ddos-attack-by-chinese-hackers/
[4] https://agwb.cag.gov.in/files/agae/circular_order/iCISAs_Study_Paper_and_Presentation.pdf
[5] https://www.sailpoint.com/identity-library/malicious-insider
[6] https://www.researchgate.net/publication/268718596_Introduction_to_identity-based_cryptography
[7] https://www.researchgate.net/figure/Private-key-generator-PKG-in-an-identity-based-encryption_fig2_360911547
[8] https://www.simplilearn.com/tutorials/cyber-security-tutorial/sha-256-algorithm#:~:text=Digital%20Signature%20Verification:%20Digital%20signatures,the%20verification%20of%20the%20signature.
[9] https://support.google.com/googleads/answer/9004655?hl=en#:~:text=SHA%2D256%20stands%20for%20Secure,will%20create%20the%20same%20hash.
[10] https://securiti.ai/glossary/secure-hash-algorithm-sha-25bit/#:~:text=Algorithm%20256%2Dbit)-,SHA%2D256%20(Secure%20Hash%20Algorithm%20256%2Dbit),data%20authentication%2C%20and%20password%20hashing.
[11] https://www.ibm.com/products/guardium-data-detection-response?utm_content=SRCWW&p1=Search&p4=43700082003997064&p5=p&p9=587000088316247 58&gad_source=1&gad_campaignid=22082087883&gbraid=0AAAAAD-_QsTSUK5ttnMRdFnWY-dJvStRn&gclid=Cj0KCQjwgIXCBhDBARIsAELC9ZhT6Q8C84VWfp3mNw9YaGPUT0WZtMHjdGr3PE-TG5ilYD5DTOsQm3oaArXqEALw_wcB&gclsrc=aw.ds
[12] https://nordlayer.com/learn/cloud-security/risks-and-threats/
[13] https://roboticsbiz.com/top-10-cloud-computing-security-algorithms/
[14] https://ieeexplore.ieee.org/document/7118923/