RESEARCH REVIEW JOURNALS
Volume-04
Issue-08
August-2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com[Peer Reviewed Journal]

Data Masking in Cloud Computing

Jhang Paul

Software Engineer, KGS Technology Group, Alpharetta, Georgia, USA

ARTICLE DETAILS

Article History
Published Online: 16 Aug 2019

Keywords
Cloud Computing, Masking data, Cryptography, Cloud security, Hiding data, Confidentiality data, Integrity data.

***Corresponding Author**
Email: jhangpaul75[at]gmail.com

ABSTRACT

Different approaches can be adopted when designing a data masking architecture to meet business needs. Although it is not advisable to store sensitive data on a cloud storage space, sometimes it may be necessary for you to have access to certain documents anywhere and on multiple computers. This paper discusses the security challenges of cloud computing and focuses on data masking and different techniques for masking data in the cloud.

1. Introduction

Cloud computing [1] has become one of most popular technology and organizations are leaning towards it across all the business domains. The main reason for adopting this technology is the ease of access to an IT resource park with almost unlimited potential, all with minimal management effort. An organization can rent the shared resources of a cloud service, becoming an infrastructure tenant rather than an owner. But there are always security risks associated with cloud computing. Indeed, cyber-attacks have shown the possibility of exploiting the proximity of sharing the same cloud service. Security issues, especially privacy and data integrity, are therefore a major concern for cloud users because their data is managed outside of their governance. In the context of government, social security and health care, these problems are even more important because they potentially concern the sensitive data of citizens and businesses. However, the benefits of the cloud can't be denied: cost savings, greater scalability, better mobility, faster deployment and instant upgrades, to name just a few. However, security risks are still the biggest obstacle to cloud adoption by organizations and businesses. The purpose of this research note is to present known and existing cryptographic mechanisms that help protect storage and data processing in cloud environments.

2. The challenges of cloud computing in the face of security

To understand the protections offered by cryptography in the cloud, we must consider security in relation to its three objectives.

Since availability is generally solved in today's cloud computing environments by non-cryptographic means, we focus only on privacy and cloud integrity [2].

Confidentiality can be provided by cryptographic encryption. Indeed, once data is encrypted, it does not "signify" itself. The key that has been used to make this encryption is the only information that can help cover the data in the clear.

Integrity can be provided by a hash function or a digital signature. Indeed, once specific original data is hashed or digitally signed, an alteration of the original data can be detected because the hashed / signed version of the data no longer corresponds to the original version.

3. Data masking in the cloud

Data masking in the cloud [3] is simply encrypting data and providing encryption keys only to users who need to be able to access them. However, encryption also has its disadvantages. Starting with the risks associated with a possible compromise of private keys. To face all its challenges and problems, it is imperative to follow certain security techniques that are categorized as follows:

3.1 Encryption

It uses algorithms to make information unreadable by transforming initial data into cryptograms. The mathematical algorithm that transforms the initial text into a pictogram, and vice versa, is called a key. See Figure 1 for a simple illustration of how encryption works.

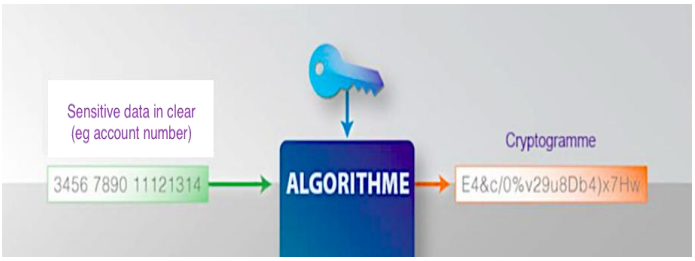


Figure 1: Encryption process

It is possible to use a single key for the encryption of information and their decryption to return to the original text format. However, in practice, it is safer to use one key for data encryption and another for decryption, best of all, a unique key pair can be used for each data field being encrypted. This practice of multiple keys adds a level of security, so that if one field was unveiled, the others would not be affected. See Figure 2 for a simple illustration of using separate keys to encrypt and decrypt data.

3.2 Tokenization

Tokenization is the process of random substitution of a value, or token (or token), used in place of real data and in

which the token is not the result of a computer calculation or a form to recover the initial value of the data. The most common form of tokenization uses a highly secure search array (called a safe) to maintain the link between the actual data and the token substitution values. This process is illustrated in Figure 3.

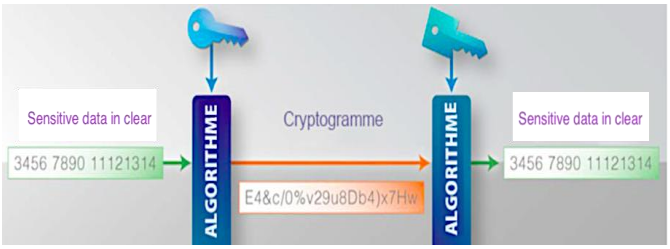


Figure 2: Encryption / decryption using different keys

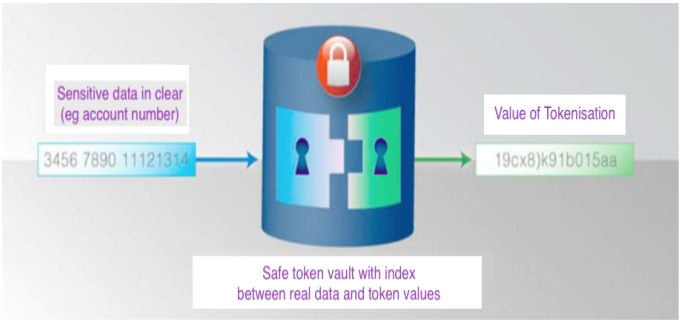


Figure 3: Data Token Process

The tokens being the result of chance, there is no connection between them. A pierced token does not help in any case to reveal others. There is no key or calculation that allows an intruder to unlock all the tokens if he drills one. Tokenization is seen as the safest way to use surrogate values for initial data in the clear. With encryption, the value of an encrypted field is reversible to recover the original data, unlike tokens.

The discipline of encrypting data in the cloud is relatively new. Cloud providers, encryption specialists, and enterprise customers are moving forward because there are few enterprise environments where applications, data, and encryption keys are in one place.

The biggest challenge is finding an effective way to mask the data so that they are not exposed in the clear in the cloud, while preserving the functionality of the SaaS application. The last point is essential as if the end user does not find the features of the SaaS application, it defeats the main generators of the company. Security professionals know they need to protect sensitive data without affecting SaaS features. This is a dilemma that can only be solved through innovation to meet all the requirements.

4. The different techniques for masking data in the cloud

The success of any project depends on the preparation and adherence to best practices. The following steps can help companies as part of a data masking project [4].

4.1 Discover the data

The first step in a data masking project is to identify the data to be hidden to protect them appropriately. This step usually accounts for 10 to 20% of the project team's work.

There is no standard approach here: it all depends on the needs of the business, the complexity of the data, and the scope studied. The deliverable for this phase should be the identification of data exposure risk, privacy issues, and how data masking will reduce the risk.

4.2 Architecture design

All data discovered is not necessarily sensitive. The classification of data is therefore a very important step in this phase. The classification of the data must be done according to the regulations and standards in force. But it's also about taking into account the functional needs to make sure that the application fed by the data studied will continue to fulfill its functions correctly once the data is hidden. A risk profile and risk tolerance model of the organization should also be established to determine what constitutes an acceptable mask level in the environment, or which rows and columns to mask.

4.3 Construction and configuration

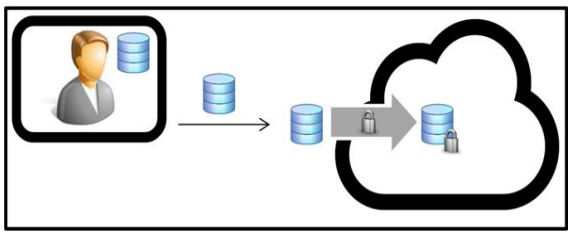
In this step, it is a question of building and putting in place the right set of configurations starting from the needs determined during the previous phase. This includes how the data masking components will be integrated and how the change management process will be initiated. Here too, it is needed to create the data masking rules and implement the necessary customizations.

4.4 Deployment

This is the final phase of the project [5]: integrating the data masking architecture into the test environment. This involves creating test databases, masquerading tasks and necessary scripts, and measuring user acceptance. This part can be subdivided into several such that:

Server-side encryption [6]

This encryption technique is the most widely used today in cloud services. As shown in the figure below, user Bob sends "unencrypted" data to be stored in the cloud service. It is the cloud service that encrypts the data before it is stored. The cloud service can also optionally apply a hash function or a digital signature to the data to be stored.



In this configuration, the management of the keys used for encryption must be server side, where two separate cases are possible.

- Non-secure storage of keys – This involves quite a bit of risks as the keys are stored in an unsecured way at the same place as the encrypted data of Bob. If Oscar can seize the encrypted data, then it is certainly Capable of grabbing the keys, and so can find Bob's clear data.

The confidentiality of the data with respect to Oscar is not assured.

The integrity of the data vis-à-vis Oscar is not ensured either if the keys used for the function of hashing or digital signature are also not protected.

- Secure storage of keys - The other possible situation is to store the keys securely, whether in the same place or not as the encrypted data of Bob. A technician in the use of an HSM3, (see example in the figure opposite), electronic device considered inviolable where it is possible to generate / store / protect keys and perform cryptographic operations with them. HSMs meet high security standards (eg Common Criteria EAL4 +) and therefore represent a high standard of security. Another technique is the use of an encrypted keystore, ie, an electronic directory containing the secret keys and protected by a password or other cryptographic mechanism, such as the PKCS # 124 standard. With both protection techniques, even if Oscar takes over the encrypted data from Bob and the HSM or keystore, he is unable to recover the keys, so Bob's clear data.

The confidentiality of the data vis-à-vis Oscar is therefore assured.

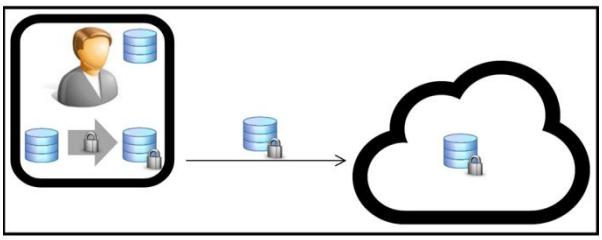
The integrity of the data vis-à-vis Oscar is also assured if the keys used for the hash function or digital signature are also protected.

Moreover, there is no confidentiality or integrity regarding the cloud service in any case. Indeed, whether the keys are in a protected environment or not, since the cryptographic operations are server side (in the cloud service), the cloud service "sees" the data in clear before processing.

Client-side encryption [7]

This other decryption technique is less widespread, although it allows the user to have full control over their data. As shown in the figure below, user Bob encrypts his data on his side and then sends his data encrypted to the cloud service for storage. Bob can also optionally apply a hash function or a digital signature to its data before sending them to the cloud service. In this configuration, the management of cryptographic keys must be done on the client side (c.- to keep the user in control of his data. The cloud service knows neither the keys nor the data in the clear: it only "sees" the encrypted data.

Note: Even if the keys are on the client side, and do not fall into the hands of our opponent type Oscar, it is not excluded to protect them locally.



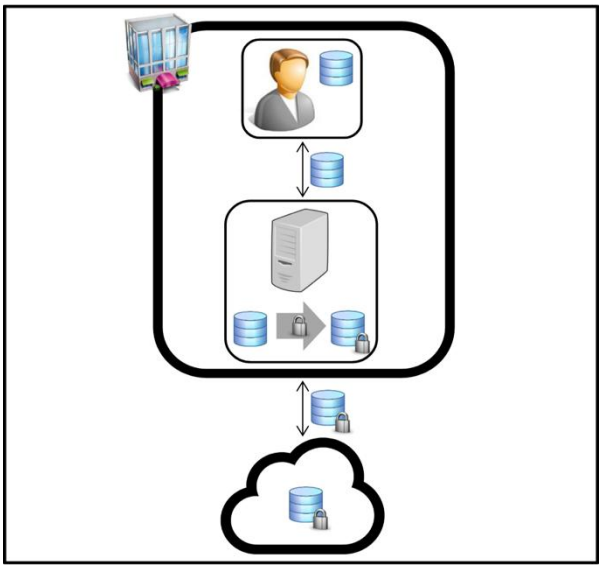
The confidentiality and integrity of the data regarding the Oscar cloud service are therefore ensured.

If the user is the only entity that manages their keys, then this can become dangerous. Indeed, if Bob loses or forgets his keys, the encrypted data are indecipherable and the data in clear are permanently lost. One solution to this problem is that the keys are managed by a TTP which has a key escrow mechanism: this allows, in certain extreme or serious circumstances, the TTP to find the keys lost or forgotten and thus to get the data in clear. Of course, it is to use sparingly. Another weak point is that the keys must be installed on all devices of the client, otherwise Bob will be unable to decipher its data from its different devices.

Cloud security gateway [8]

Finally, another technique to protect the data stored in the cloud is the use of a Cloud Security Gateway (CSG). This technique has already been the subject of several publications. A CSG is a tool that is located between a user and a cloud service, usually within the user's enterprise in the professional setting. It is a trusted intermediary from the security point of view, commonly in the form of a proxy, which encrypts / decrypts data that pass through it. As shown in the figure below, its use is fairly straightforward: user Bob sends his data in clear to the CSG, which encrypts this data and sends it to the cloud service.

Note: Even if the keys are on the client side, and do not fall into the hands of our opponent type Oscar, it is not excluded to protect them locally.



5. Conclusion

Cloud [10] SaaS applications are becoming increasingly popular with customers in almost every industry and around the world. It is rare for a client company to have any constraints on the strict protection of sensitive data. Sometimes, real innovations are needed to create a masking solution that complements the SaaS application so that the company can reap the benefits of SaaS and meet all specifications. All of this is possible, companies can finally stop giving up the cloud and to its many benefits. However, there are a multitude of cryptographic mechanisms [9] that can hide data in cloud environments. Depending on the level of security required,

different techniques can be deployed; classic encryption, encryption client and server side or cloud security gateway for processing. Each end of section brings a conclusion on the presented solutions. No solution is yet perfect, and academic research has yet to perfect all these techniques to make them more foolproof. In particular, the development and practical implementation of homomorphic encryption is expected at the turn, because this technique can bring security and ease of calculation hardly equal. Finally, even if no solution is yet perfect, most are possible to protect the data in the cloud, provided that the cryptographic keys used are under the control of the user / organization.

References

[1]. Jathanna, Rohan & JAGLI, DHANAMMA. (2017). Cloud Computing and Security Issues. International Journal of Engineering Research and Applications. 07. 31-38. 10.9790/9622-0706053138.

[2]. Abdullatif, Firas &Zuhiar, Maan. (2017). Cloud Security Issues and Challenges: Important Points to Move towards Cloud Storage. International Journal of Science and Research (IJSR). 6. 6-391. 10.21275/ART20176320.

[3]. K.Sharmila S. Borgia Anne Catherine Sreeja V.S, "A comprehensive Study of Data Masking Techniques on cloud", International Journal of Pure and Applied Mathematics Volume 119 No. 15 2018, 3719-3727.

[4]. Yesilyurt, Murat &Yalman, Yıldıray. (2016). New approach for ensuring cloud computing security: using data hiding methods. Sādhanā. 41. 1-10. 10.1007/s12046-016-0558-8.

[5]. Gupta, Diksha & Chakraborty, ParthaSarathi& Rajput, Pragya. (2015). Cloud Security Using Encryption Techniques. International Journal of Advanced Research in Computer Science and Software Engineering. 5. 5

[6]. C. Gopinaath and C. Kiruthika, "A Server Side Encryption for Cloud Storage with Federation Sharing in Hybrid Cloud Environment," *2017 International Conference on Technical Advancements in Computers and Communications (ICTACC)*, Melmaurvathur, 2017, pp. 128-131. doi: 10.1109/ICTACC.2017.41

[7]. Sugumar, Ramalingam, and K. Raja. "A Study on Enhancing Data Security in Cloud Computing Environment." performance computing 6.3 (2018).

[8]. Simanta Shekhar Sarmah, Cloud Migration- Risks and Solutions, Science and Technology, Vol. 9 No. 1, 2019, pp. 7-11. doi: 10.5923/j.scit.20190901.02.

[9]. Shah, Nidhi, and Digvijay Mahida. "Data Security in Cloud Computing: A Comprehensive Survey." (2018).

[10]. K. Sharmila and Dr.S.A.Vethamanickam, " MRK-SVM: An Effective Technique for Big Data In Health Care Sector",International Journal of Scientific & Engineering Research, Volume 7, Issue 6, June-2016,ISSN 2229-551

