Vol. 10(1), pp. 1-8, January-June, 2024
DOI: 10.5897/IJCER2023.0308
Article Number: DFC69D271971
ISSN: 2141-6494

http://www.academicjournals.org/IJCER

ACADEMIC JOURNALS
expand your knowledge

International Journal of Computer Engineering Research

Review

A practical guideline for implementing an extra security layer on an intra-cloud private MongoDB Database using encryption

Kevin Tshimanga[1*], Patrick Mukala[1,2] and Godwill Ilunga[1]

[1]School of Computer Science, Faculty of Engineering and Computer Science, University of Wollongong in Dubai, Dubai, United Arab Emirates.
[2]Department of Mathematics and Computer Science, Faculty of Science, Université Pedagogique Nationale, Kinshasa, Democratic Republic of the Congo.

Received 7 September, 2023; Accepted 28 December, 2023

Cloud computing offers different deployment and distribution models for the outsourcing of the company's Information Technology (IT) infrastructure (applications and data). The wide adoption of this concept does not exclude permanent and unforeseen dangers. The company loses its guarantee of the confidentiality of its most sensitive information, and many legal and security questions remain unanswered. A large number of threats from different backgrounds flock, especially against the data. The confidentiality of sensitive database information stored on the cloud provider server is always a big issue for cloud customers. Because it is no longer just attackers or other cloud users the company is looking for to preserve the confidentiality of its sensitive information, it is also and above all the supplier cloud itself. To this end, cloud providers have deployed security mechanisms designed to protect user databases from external threats. However, these measures do not seem sufficient. In the absence of trust, security does not exist. In this thesis, we set out to define the main database security factors. We have made relevant arguments on the necessity of the client-side encryption model and evaluated its performance in an environment real cloud. In order to carry out these measurements, we propose a basic security model of data based on client-side encryption. Our proposal guarantees the confidentiality of data, thanks to the management of the encryption/decryption keys assigned to the client cloud.

Key words: Cloud computing, database security, data encryption, confidentiality, integrity, trust.

INTRODUCTION

Today, more and more companies are computerizing their structures. This allows them to better manage the flow of information on a daily basis. However, when we talk about the computerization of any structure, from a physical point of view, we immediately see a computer architecture made up of several computer devices such as servers and local networks. However, in most cases, we ignore the software aspects that should not be

*Corresponding author. E-mail: kevin.tshimanga@outlook.fr.

neglected, such as databases, which will become an excellent storage space for these data streams.

"6% of businesses that experience data loss survive, 43% shut down permanently, and 51% collapse within two years" (Date, 2004).

The company's IT department does not always have all the expertise and skills required to set up and maintain an efficient and secure storage architecture. Cloud computing has come at the right time as an option to solve this problem. This simple concept essentially suggests that the company's information and services should be outsourced to large warehouses called data centers. Experts in the field manage these physical sites; Amazon, Google and Microsoft presented their solutions and provided support for data storage. Lets explain here that migrating the company's IT resources to cloud computing can significantly reduce the operating costs of the company that subscribes to the product.

However, companies that choose to migrate their data to cloud computing are somewhat losing physical control over their information. It thus accepts the principle of the Service Level Agreement (SLA) which is defined as a contract concluded between a customer and its cloud provider.

Any company that gets involved is legitimately entitled to seek answers to the following questions:

(1) Does the cloud provider have access to my data? Is it exploiting them outside of my control?
(2) What would happen if the infrastructure were ever to be compromised?
(3) What are the major challenges? Possible approaches? The significant challenges of implementing a customer-based model?

These are all questions that triggered the problems of this work, namely, how to place the user of a cloud service at the center of the security of his data or at least, how to give him more credit for legitimacy. Protection of the most sensitive data in its collection.

The scandal linked to the National Security Agency's (NSA) PRISM[1] global surveillance program in October 2013, or more recently, the one involving "Cambridge Analytica - Facebook"[2] highlight our questions and provide concrete but worrying answers.

In this paper, we delve deeper into the literature review to explore how cloud protection systems are implemented and what kind of network security mechanisms can be implemented to maintain security policies in the networks they safeguard. Finally, in the results and discussions, we introduce a client-based security model in response to the problem described earlier.

LITERATURE REVIEW

"Cloud computing is a model used to achieve ubiquitous, convenient, and on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services)" (Kumar et al., 2017b).

Cloud computing has changed the perception of information technology (IT) capital and operating expenses and has changed the way infrastructure is designed in today's computing age. Cloud computing has also dramatically reduced the costs of starting new businesses and affected the way we store and process data. Companies can choose cloud computing for many key business goals, such as lowering operating costs, reducing complexity, immediate access to resources, easy expansion, and lowering barriers to innovation (Kumar et al., 2017a, b).

Cloud protection system

Cloud computing continues to transform the way businesses use, store and share data, applications and workloads. It also introduced a host of new security threats and challenges. With so much data entering the cloud, and public cloud services in particular, these resources become natural targets for bad actors.

"Cloud computing and web services run on a network fabric, so they are vulnerable to all types of network attacks" (Alliance, 2020).

Clearly, security concerns have played the biggest role in hampering cloud computing. There is no doubt that using our data, with someone else's processor to run our software on someone else's hard drive can seem intimidating to many people, especially when you do not know what our data can be used for when we do not keep our eyes on it all the time.

In order to keep organizations abreast of cloud security issues so that they can make informed decisions about cloud adoption strategies, the Cloud Security Alliance (CSA)[3] has created the latest version of the main threats to the cloud: Egregious Eleven (Alliance, 2020). The report was released in September and lists the top cloud threats that occurred in 2019. To identify key issues, CSA interviewed industry experts to gather professional

[1] PRISM is a code name for a US program for the mass surveillance of worldwide internet communications

[2] The Facebook-Cambridge Analytica scandal or the Facebook-Cambridge Analytica data breach refers to the personal data of 87 million Facebook users that Cambridge Analytica (CA) began exploiting in early 2014.

[3] Cloud Security Alliance is a non-profit organization whose mission is to promote the use of best practices to ensure security in cloud computing and to provide training in the uses of cloud computing to help secure all other forms of computing.

opinions on the biggest security issues in the world.

Network security mechanism in the cloud

The purpose of the network security mechanism is to ensure the security of hosts and applications on the network to which it belongs. We are particularly interested in these network security mechanisms:

(1) Network access control, performed by a firewall or firewall.
(2) Intrusion detection, performed by an intrusion detection system (IDS).
(3) Encryption.

These mechanisms are designed to implement security policies within the networks they protect.

Virtual firewalls

A firewall is a tool allowing to control the traffic circulating between the inside and the outside of a security perimeter (Mè, 2006) which constitutes the limit between the network which one considers to be secure or which one wishes to protect and the rest of the Internet. Firewalls can provide a variety of services, but we are interested in the basic services, which is network-level access control. The main function of network access control is the filtering of data packets, which is operated by filter rules that prohibit or allow certain types of traffic. These rules are established according to the parameters of the protocol header located in the data packet. They are applied to each packet passing through the firewall, in a specific order of priority in the application of the rules. Most firewalls are generally stateful, that is, before allowing the packet, they will additionally check whether the packet belongs to the current connection (such as a Transmission Control Protocol “TCP” connection). Therefore, they check whether each packet allowed by the rule initiates a new connection or is part of an existing connection. In addition, different levels of inspection can be performed on the packaging. The most common level is Deep Packet Inspection (DPI), which thoroughly examines the contents of a packet to allow or deny access. Web Application firewalls (WAF) dedicated to the protection of web servers are also increasingly common (Mè, 2006).

Intrusion detection system

According to Powell and Stroud (2003), intrusion detection involves all the practices and mechanisms used to detect errors that may cause security breaches and/or detect attacks. Still according to Powell and Stroud (2003), IDS is an implementation of intrusion detection mechanisms and practices. Therefore, the role of IDS is to detect and/or prevent (in this case we call an intrusion prevention system) attacks that occur in a system or network. They can be deployed on a monitored host, called a host-based intrusion detection system (HIDS), or on a monitored network, called a network-based intrusion detection system (NIDS).

Encryption

Encryption is the process of converting information into a number or code so that it cannot be read by anyone except those who have a cipher text key (Deepika and Soni, 2015). Encrypted text or coded text is called encrypted data. This can be achieved using different types of encryption methods such as Message-Digest algorithm “MD5 SQL” encryption, asymmetric public key encryption, Advanced Encryption Standard (AES), cryptography, etc. Encryption is the foundation of the solution model proposed in this article, which enables an effective security mechanism to be put in place to preserve the confidentiality of sensitive data of cloud clients.

RESULTS AND DISCUSSION

At the beginning of this work, the first stones were laid, which formed the common thread of this study. A problem based on concerns that anyone wishing to migrate their data to a public cloud infrastructure is entitled to have, was identified.

(1) The question was whether the customer’s data could be at risk on the servers of a public cloud provider.
(2) In addition, if a cloud provider could have access to its customers’ data; even the most sensitive data.
(3) Finally, be sure that if the cloud provider’s infrastructure were to be compromised, the measures put in place could ensure that customer data will be unusable for the attacker.

Modeling the solution

Because pictures speak better than words, we provide a set of graphics that should better describe our solution to the original problem, which was to provide cloud customers with more privacy over their data stored in the cloud.

Data encryption is not new in terms of the security of data flows within an information system.
Encryption opportunities are available in several stages. The critical questions are:

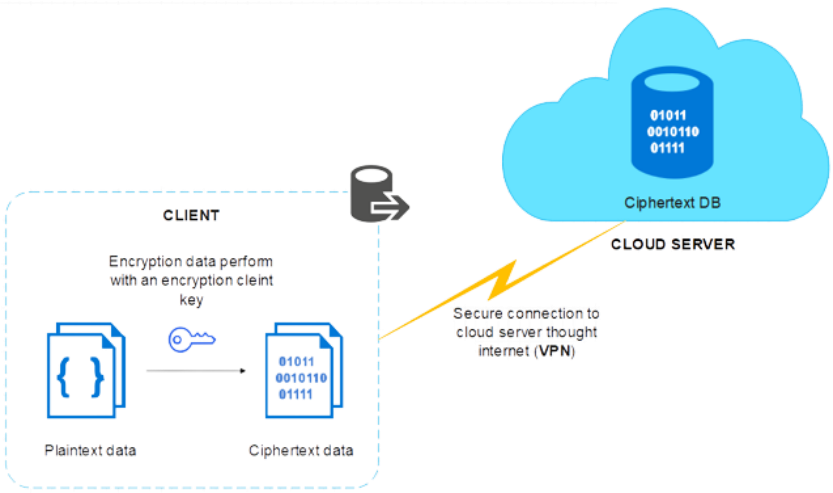

Figure 1. Client-side encryption of the DB before storing it on a cloud server.

(1) Where should the data be visible in the clear?
(2) Who should own the decryption keys?

Figure 1 is a representation of the solution model we propose to answer these questions.

(1) The data should only be clearly visible to the owner of the data, that is, the cloud client. Therefore, data encryption is the best solution to achieve this kind of data privacy.
(2) The management of encryption/decryption keys is the responsibility of the data owner (cloud client). Hence, this means that the cloud provider will only be responsible for the storage of the encrypted data by its customers.
(3) The master encryption key will be backed up on the client-side; which will allow the customer to access their data stored on the cloud servers, they will use the key in their possession to decrypt the information.

After storing its encrypted data on the cloud server, the cloud client will want to access it eventually; at this point, the entire database will no longer be returned to the client; but there are only very specific records that customers can easily find through queries. The procedure to be followed is illustrated in Figure 2.

(1) When the client tries to access the encrypted data on the remote database, since the server does not have any decryption key, the server will send encrypted data to the client, so it is impossible to decrypt the data before sending it back to the customer as it is done in the existing cloud security system.
(2) This perfectly answers the question of whether the supplier can have access to its customers' data. Indeed, there can be access but without any possibility of using them since they are encrypted and therefore unusable without the decryption key.

Algorithmic approach

Encryption method

Generate Master Encryption Key on Client Side
(1) Master_Key
Set up a local KMS Provider to store Generated Master_Key
(2) KMS_Provider --> (Master_Key)
Generate Secondary Encryption Key stored on Cloud Server Side
(3) Master_Key --> Second_Key
Encrypt Secondary Key Using Master Key
(4) Encrypt (Second_Key) using Master_Key
Encrypt Personal Cloud Data Using Secondary Key

Decryption method

(1) Decrypt Secondary Encryption Key Using Master Key
(2) Decrypt Cloud Data using Decrypted Secondary Encryption Key

Client-side field-level encryption

Client-side field-level encryption (CSFLE) provides an additional layer of security for the most sensitive data. Using a MongoDB driver, CSFLE encrypts certain fields that the client specifies, ensuring that they are never transmitted unencrypted, nor seen unencrypted by the MongoDB server. [MongoDB,]

It also means that sensitive information cannot be obtained from the database server. Without access to a specific key, data cannot be decrypted and exposed. Reading data directly from disk, even with DBA or root credentials, will also be impossible because the data is

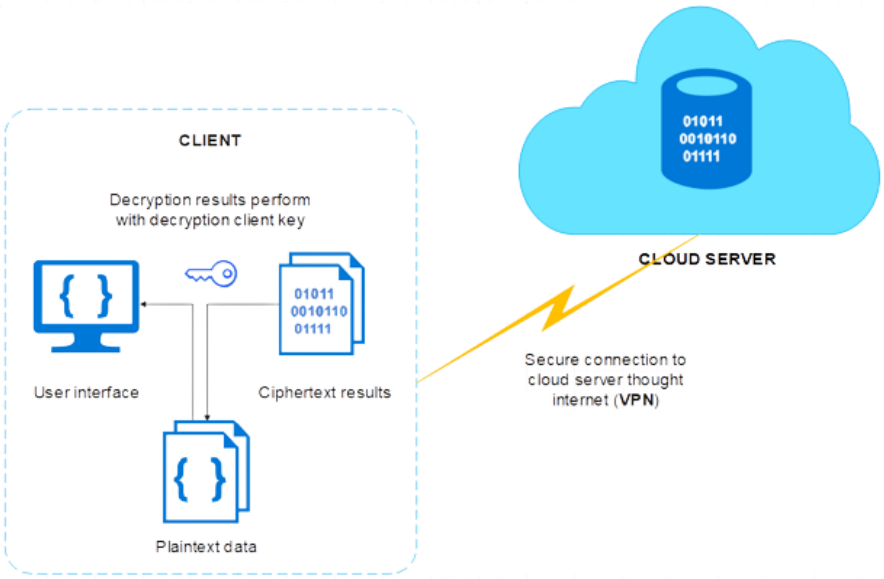

Figure 2. Decryption of the encrypted result of a request to the server.

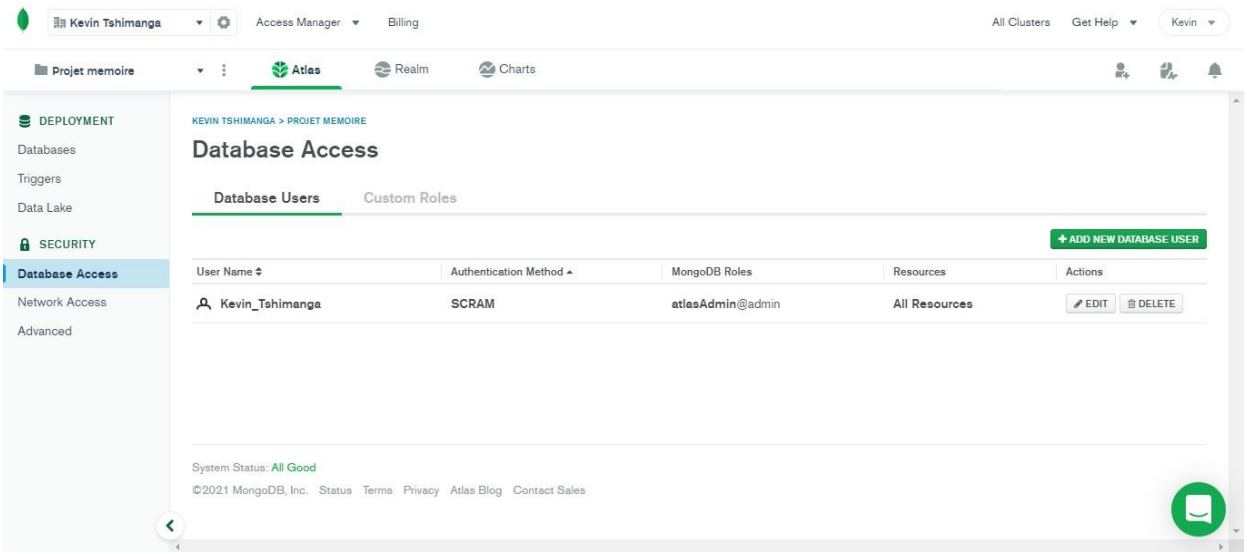

Figure 3. Creation of a user account (Administrator). (captured from MongoDB Interface)

stored in an encrypted state.

Testing the CSFLE security model with MongoDB

Here, is mainly taken from our testing of the solution model offered in the MongoDB Atlas cloud environment (Figures 3 to 8).

Conclusion

Cloud computing is undoubtedly a concept that goes beyond how to provide services to different customers, whether for businesses or the public. We believe this idea is technically noble, it can save a lot of cost, while removing the burden of managing complex IT infrastructure for customers.

However, there are still many security management issues; we have identified them in our issues. These issues relate to the management of data security by cloud clients; companies cannot be guaranteed the confidentiality of their most sensitive data when it is hosted on their provider's cloud servers. In addition, scandals such as the PRISM incident of giants like

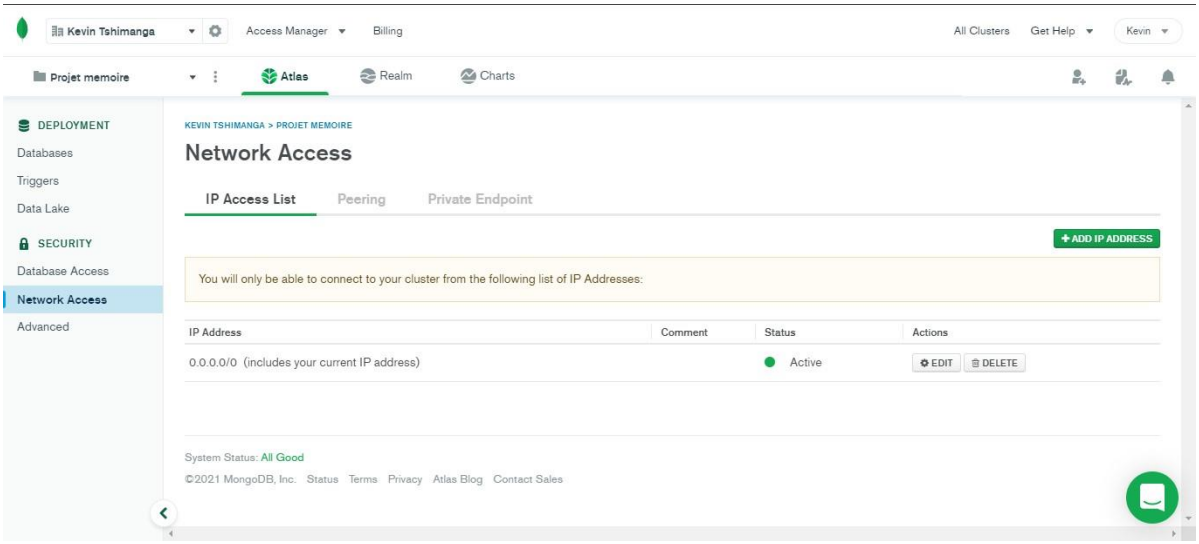

Figure 4. Configuration of authorized IP addresses. (captured from MongoDB Interface).

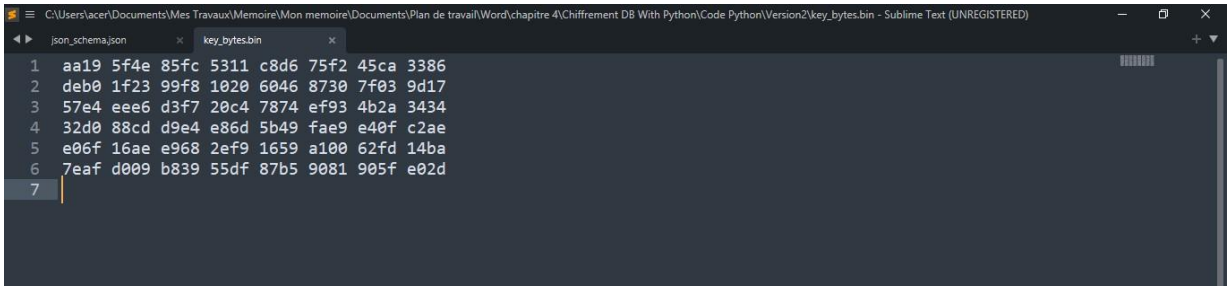

Figure 5. Generating the master encryption key. (captured from a text editor).

```
C:\Users\acer\Documents\Mes Travaux\Memoire\Mon memoire\Documents\Plan de travail\Word\chapitre 4\Chiffrement DB With Pyth
on\Code Python\Version2>python csfle_main1.py
RESULTATS DECRYPTES() TROUVES:
 * {'_id': ObjectId('6135f1673c077286435e50ac'), 'full_name': 'Kevin Tshimanga Kabongo', 'ssn': '123-12-1234', 'groupeSang
uin': 'A+'}
 * {'_id': ObjectId('6135f16b3c077286435e50ad'), 'full_name': 'Jonathan Lawamu Kasimu', 'ssn': '523-22-1234', 'groupeSangu
in': 'B+'}
 * {'_id': ObjectId('6135f16c3c077286435e50ae'), 'full_name': 'Abigaïl Bilonda Mukendi', 'ssn': '452-33-1234', 'groupeSang
uin': 'O-'}
RESULTATS CRYPTES() TROUVES:
 * {'_id': ObjectId('6135f1673c077286435e50ac'), 'full_name': 'Kevin Tshimanga Kabongo', 'ssn': Binary(b'\x02\x10;\xe4jnTM
\xf3\xad)\xe0\x9a\x18o.\xdd\x02\xeeM\xe47\xcfY\xb6\xd6\x1eX\xc9;7\x7f[\xdf\\\x9f\x05L:\xf4\x94\xf0c\xefsgj\x9a\x01\x91\x12
=\xfd\x90k;\xebS\x80:tN\x85\x10\xd7\xb6\xb4@\xe4]\xa1\xd8rQh\xeb\xd7\x91\xb0\xb0\xc3[q\xcc\x01\x81\xd5\x92j\xa6\xd3\xf2\xe
2\x07R\x16\xe4\x9e', 6), 'groupeSanguin': Binary(b'\x02\x10;\xe4jnTM\xf3\xad)\xe0\x9a\x18o.\xdd\x02\x89,\xebS1<\x11\xf3\xd
0W\xcd\x95\xd0\x87\xce\x08\x02\xcc5\xbe\x021\x81\xa2|g8\xcd(\xc5\x9dy4\x98&6&)\x90(\x8a\xcc\x8cg\x80y\xf1\x8fI\x86\xd9\x02
\xf60\xcbP\xe0\xd7F\xf5})2\xff', 6)}
 * {'_id': ObjectId('6135f16b3c077286435e50ad'), 'full_name': 'Jonathan Lawamu Kasimu', 'ssn': Binary(b"\x02\x10;\xe4jnTM\
xf3\xad)\xe0\x9a\x18o.\xdd\x02\xad\xef\xdc\x8e}I\n\x81\xec\x90\xe1\xa1\xce\x9cF\x87\xa0\xb5\xf06D\x1cO\xa9\xed\xde\x8eE\x9
a\xf7\xc1M\xd5\\\x12\x81\x15O\xa4\xf91\x94\x9d\x0b'\xcd32?m\xf1\xc1s\xaf\xf5\xeaZ\x96\xfa\x1b>/\xde2\xbc3*\x13\x14i1K\xf4j
\xd1\x92\x93)\xb4\xbd", 6), 'groupeSanguin': Binary(b'\x02\x10;\xe4jnTM\xf3\xad)\xe0\x9a\x18o.\xdd\x02\x93u\x85\xcc\xfcPF\
xab ~\xcd\xdfc\xc3\xcf\\\xf8j\x1c\x98\xee7\xcd\xa7\xd7\x9bjw\x1f\xf0\x9a\xbb7\x9b&\xa1\x9aPr\xd1\xde\x17\xcc\xc9\x01N%\xc2
Km\x98\xdf\x12?\xa4\xc6\xce\x0ca\xc4\x02\xb5\x84}', 6)}
 * {'_id': ObjectId('6135f16c3c077286435e50ae'), 'full_name': 'Abigaïl Bilonda Mukendi', 'ssn': Binary(b'\x02\x10;\xe4jnTM
\xf3\xad)\xe0\x9a\x18o.\xdd\x02\xcd\xc3\xeaS.@\x8c\xea\xbc\xf3\xfa&\xc7\xfe5:^z\xe69\x1a\xcfh\x96\xebs\xc2u\xado\xd7\xcb1\
x9b\x88>\x89\x85#\xe3;\xa4\xa7 \xf1XUj\x96\x8f\xf7\x8a6n\x1f\xaa\xef\x11\xeeQnx\x99E\x86;B\xa7\x9bU\xed(\x96\x0c\xe1\x9f\r
:\xae\x01', 6), 'groupeSanguin': Binary(b'\x02\x10;\xe4jnTM\xf3\xad)\xe0\x9a\x18o.\xdd\x02\xf8\xb8*\xeeC}r\x88\xe7\x1c\x97
B\xcav\xde\x19\xe6WS^\xf1\xc7\xf0\x96\xa8\x93rK\x0c\xbc8\x8ed\xa7X\x95\xeb\x97*J\xcf\xe1\xcaR\xc4\xd1\x82e\xca\x07\x083\x1
bb\xeb|]\x8eg\xb2\xd9 \xa0\x99', 6)}
```

Figure 6. Client-side data access with master key.(captured from a command prompt).

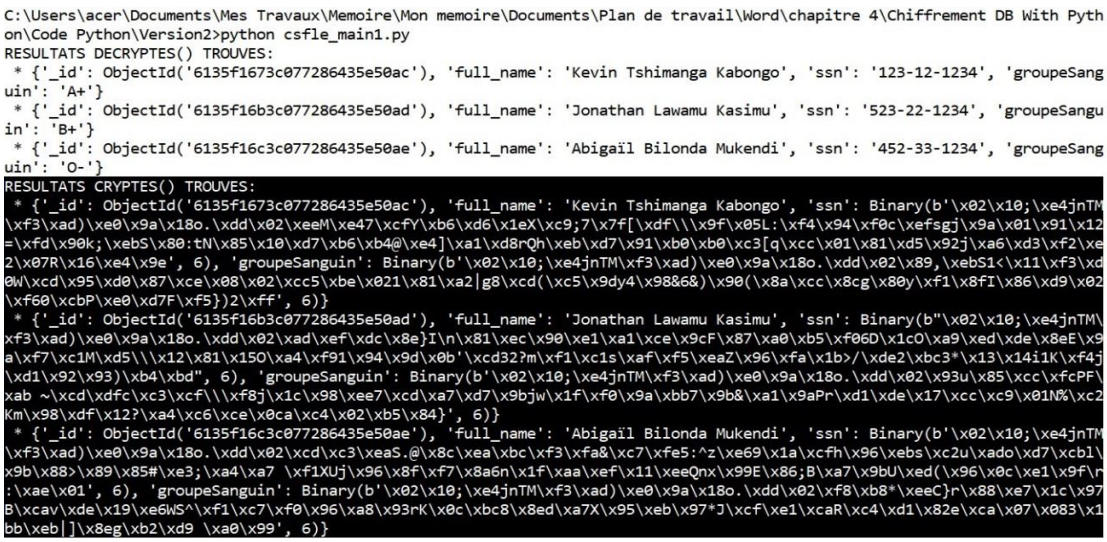

Figure 7. Accessing data from the client application without the master key. (captured from a command prompt.

Figure 8. Access to data from the server. (captured from MongoDB Interface).

Google and Facebook leaking customer information under legal constraints have further exacerbated customer concerns about the use of these services.

MAIN CONTRIBUTIONS

Faced with this observation, we provide a relevant response to the need to introduce a client-based security model: client-side data encryption. We offer a model for ensuring the confidentiality and integrity of the most sensitive data based on the cloud client. The proposed model guarantees the confidentiality of sensitive data against highly privileged users, against attackers trying to steal information in the cloud, and even against the cloud provider itself; through data encryption, the management

of encryption/decryption keys of which is assigned to the customer.

Regarding the performance test of the proposed model, two scripts were written in python, which can be consulted in the Appendix of this work for future work. The latest CSFLE technology developed by MongoDB for client-side field-level encryption has allowed us to test our security model in a real cloud environment (MongoDB) and achieve very satisfactory results.

LIMITATIONS AND FUTURE WORK

Based on our performance tests in a real cloud environment (MongoDB), we can only be satisfied with the results obtained. However, one of the points that can be raised and which could be the starting point for future work is to consider storing the encryption/decryption keys in a secure environment such as Azure KMS, Google KMS, etc.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Alliance CS (2020). Top threats to cloud computing Egregious eleven deep dive. Cloud Security Alliance 1(0).

Date JC (2004). Une introduction aux systèmes de bases de données. Boston: Pearson/Addison Wesley.

Deepika SN (2015). Database security Threats and security techniques. International Journal of Information Sciences and Techniques, 5.

Kumar V, Chaisiri S, Ko R (2017a). A data-centric view of cloud security. In Data Security in Cloud Computing. Institution of Engineering and Technology pp. 1–17.

Kumar V, Chaisiri S, Ko R (2017b). Data Security in Cloud Computing. Institution of Engineering and Technology.

MongoDB. Client-side field level encryption guide. https://docs.mongodb.com/drivers/security/client-side-field-level-encryption-guide/Consulté: 2021-09-06.

Mè L (2006). Sécurité des systèmes information.

Powell D, Stroud R (2003). Conceptual model and architecture of maftia. Technical Report Series-University of Newcastle upon Tyne Computing Science.