

MULTI-CLOUD TOKENIZATION MODEL FOR SECURED DATA ACCESS

Ajay Kumar Lebaka¹, Ratna Kumar Jala²

¹ M. Tech Student, Dept. of Computer Science Engineering, BITS Vizag, Andhra Pradesh, India

² Head of the Dept. of Computer Science Engineering, BITS Vizag, Andhra Pradesh, India

ABSTRACT

Now-a-days, many organizations are migrating quickly towards cloud computing environment because it offers various services which includes IaaS (Infrastructure-as-a-Service), SaaS (Software-as-a-Service), PaaS (Platform-as-a-Service), Storage, Database, Security, Process, Information, Application, Integration, Management, Testing-as-a-Service. These services are offered by cloud providers on pay-as-you-go basis. Companies no need to worry in maintaining storage disks and servers if they use cloud service's platform to run their applications. Even the applications, services, or resources made available to users on demand via the internet from a cloud computing provider's server. As Cloud computing systems are internet based, there are few concerns with the safety and privacy in managing sensitive data stored remotely. There is a possibility of data theft by hackers in such a case the entire sensitive data can be misused. To overcome this situation, cloud servers adapted various encryption policies to secure the data and this makes the process little bit slow when storing and retrieving the data. So, in this paper we proposed Multi-Cloud Tokenization Model which uses multi-cloud database servers for storing sensitive data along with uniquely generated tokens in one cloud database server and connecting these tokens to main user account and storing into another cloud database server. These tokens are generated in the database and there is no real meaning or mathematical relationship with the original data. But each real data is mapped to token that is uniquely generated in database either randomly or through any means like GUID or UUID (Globally or Universally Unique Identifier) and stored in token vaults, and only main identifier is mapped to generated tokens and stored in cloud database server. As we are not using any encryption policies the data processing and storage speed is comparably good than any of the previous methods. Even this tokenization can be offered as a cloud service for companies and users too. As per our study, there exists tokenization systems from long back but a formal study on this type of multi-cloud model has not been done, yet. So, we can say this process is very efficient and reliable to use after some research and thorough testing.

Keyword:- Tokenization, Cloud Computing, Multi-Cloud, Token Vault, Tokenization-as-a-Service, Data Security, De-tokenization.

1. INTRODUCTION ON TOKENIZATION AND CLOUD COMPUTING

Tokenization is the process of substituting sensitive piece of data with an equivalent non-sensitive piece of data, referred to as a token, that has no meaningful value when system is compromised, so that it can store the original data in a secure cloud data vault. Tokens acts as a reference to the original data, but cannot be decrypted. That's because, unlike encryption, there is no mathematical relationship between the token and its original data to transform the sensitive information into the token. The transformation from original data to a token uses methods like random generators, unique identifiers or keys in database as generating identifiers in database is quite fast when compared to any other generation like encryption or hashing. Similarly, De-tokenization is, of course, the reverse process, when the token is just swapped for the original number. De-tokenization can only be done by the original tokenization system and there is no other way to obtain the original number from just the token. Tokens can be of single use for transactions involved with payment gateway services and other system logins as they are not retained, or multi-use for non-financial transactions and are stored in a database for recurring transactions on same data retrievals.

According to the official **NIST** (National Institute of Standards and Technology) definition [3], “Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction”. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models. The schematic representation is shown in Fig-1.

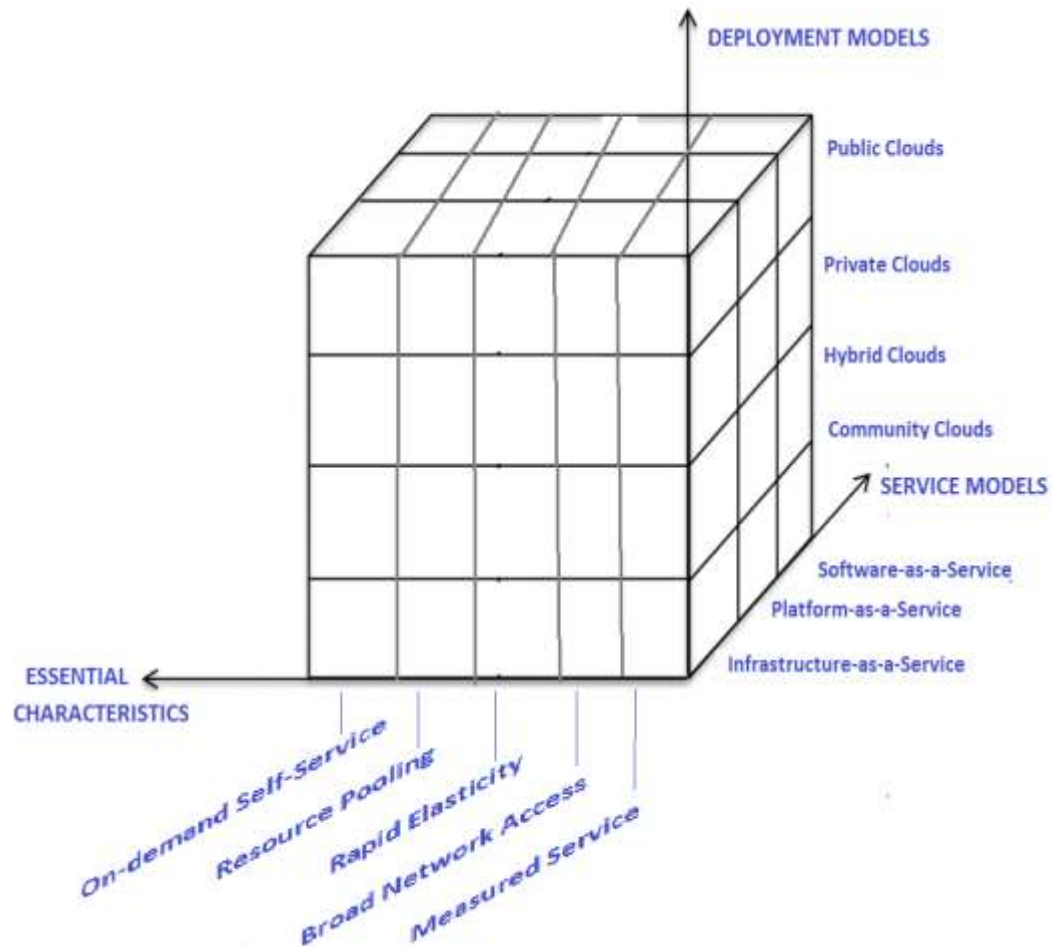


Fig- 1: The NIST Cloud Definition

Though there are several types of services provided by the cloud to the people and companies, there is a lack of proper security controlling mechanisms and policies in protecting user’s data and do not offer much customizability, for this reason many clients are not willing to move to cloud computing environment. But there are cloud providers who are investing billions in research to give better security to the user’s data by providing latest infrastructure like Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), OpenStack, SAP HANA Cloud Platform, IBM Bluemix, Oracle Cloud and many more.

Multi-cloud describes an environment that relies on multiple clouds — such as OpenStack, Microsoft Azure or AWS. For instance, you may be running a workload that requires large pools of storage and networking resources on a private cloud, such as OpenStack. At the same time, you may have a workload that needs to scale up or scale down quickly on a public cloud, such as Microsoft Azure or AWS. Each workload is running on the ideal cloud, but now you have multiple clouds to manage [2].

So, securing client data is an important criterion for good quality of customized and secured services. cloud computing faces the challenge of security threats for number of reasons. Firstly, adapting the traditional cryptographic approach for the aim of data security in cloud computing is a threat as the data are stored in remote location and users do not have any control on it. So, it requires a data verification approach and it has no explicit knowledge and correctness of the whole data as it is in third party's location. Secondly, the data are stored in third-party data warehouse and it can be frequently updated by the user, for backing up and deletion or insertion and other data alterations and in this process the other users might be affected as for every modifications encryption and decryption takes place if exists and it delays the process.

The following aspects are summarized as our contributions on multi-cloud tokenization model proposed and its benefits over cloud data encryption.

2. CLOUD TOKENIZATION AND BENEFITS OVER CLOUD ENCRYPTION

Here are few common reasons to use tokenization instead of encryption in cloud environment:

1. Data processing speed is high,
2. Reduced Cost,
3. Scalability,
4. Productivity
5. Performance improvement
6. Reliability,
7. Data Security.

3. SCHEMATIC SYSTEM ARCHITECTURE FOR PROPOSED MODEL

In general, most of the organizations uses web services from one cloud service provider and all applications, files structures, databases will reside in the same cloud environment that is diagrammatically represented in Fig- 2. This causes dangerous security threat to the user's data even though encryption policies implemented in this system because if entire system is compromised then all data is lost and can be used by the unauthenticated users. So, we proposed a new method of storing and accessing data using multi-cloud tokenization model that is schematically represented in Fig- 3.

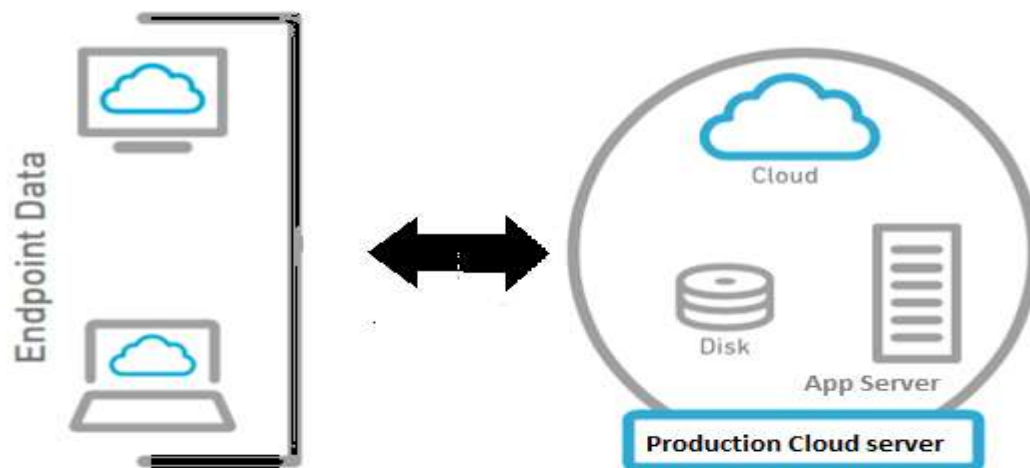


Fig- 2: Existing Cloud Environment for applications

In the proposed model, there are minimum two cloud servers one for generating tokens with the help of database server and application server and storing the index and token in database and other for storing actual data with an index that is produced with the token vault in the first server. These mappings are not related to one database

and the token vault generated here has no meaning and there is no mathematical relationship exists with original data. It is just a unique identifier generated with the help of databases and no encryption is required here.

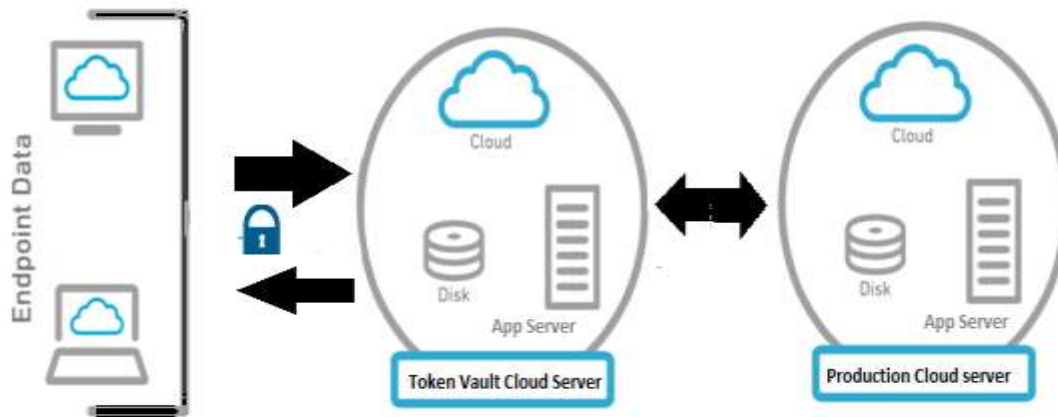


Figure 3: Proposed Multi-Cloud Tokenization Model

Our new proposed system uses the access code generated by taking the user given input and sent the same to the user by email. In the new login process, user first give the user given access code, and then user has to give registered email address given at the time of generating unique access code. We will authenticate the user basing on user given access code & registered email address, then only we will proceed further into the cloud services provided by the Cloud Service Provider (CSP).

4. CONCLUSION

Cloud computing is amazing technology with increasing demand and popularity and industry giants like Amazon, IBM, Microsoft, and Google, have been promoting cloud computing by adapting latest equipment and gaining lots of costumers. But the rest of the public that are still doing research on the topic are still ambiguous to adapt this due to security threats and confusion over the security and are afraid to migrate to the cloud. But this is faster, cheaper, and easy to use technology and users should move to this rapidly moving technology.

In this paper, we gave a solution to secure data by using multi-cloud tokenization model without the need of encryption and not compromising the speed of storing or retrieving the data. So, this is cost-effective and efficient model for securing the data that is in the cloud. But, one must do lots of research and testing before adapting to this model from the existing system.

In the future, we would like to extend our cloud tokenization model for providing tokens to the third-party clients for their input data by establishing Tokenization-as-a-Service and extend to make use of these token vaults to map various types of data to single user accounts on multiple servers by adapting Single Sign On (SSO) policies. Part of our work is still in the testing process and anyone can adapt this approach with proper testing and research.

5. ACKNOWLEDGEMENT



We would like to thank our family members and friends for their support and valuable suggestions. I would also like to thank Mr. K.S.N. Murthy, P.G. Coordinator, Dept. of Computer Science and Technology, BITS College, Vizag and to my friend and colleague Mr. Kandala Venkata Jagannadham for their assistance.

6. REFERENCES

- [1]. Hassanein, Hossam & Elragal, Ahmed. (2014). "Business Intelligence in Cloud Computing: A Tokenization Approach"

- [2]. Rackspace cloud library, online at <https://www.rackspace.com/en-in/cloud/multi-cloud>
- [3]. "The NIST Definition of Cloud Computing", NIST Special Publication 800-145 by Peter Mell and Timothy Grance.
- [4]. H. Graupner, K. Torkura, P. Berger, C. Meinel and M. Schnjakin, "Secure access control for multi-cloud resources," 2015 IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops), Clearwater Beach, FL, 2015, pp. 722-729.
- [5]. S. Díaz-Santiago, L. M. Rodriguez-Henriquez and D. Chakraborty, "A cryptographic study of tokenization systems," 2014 11th International Conference on Security and Cryptography (SECRYPT), Vienna, 2014, pp. 1-6.
- [6]. Z. C. Nxumalo, P. Tarwireyi and M. O. Adigun, "Towards privacy with tokenization as a service," 2014 IEEE 6th International Conference on Adaptive Science & Technology (ICAST), Ota, 2014, pp. 1-6.
- [7]. Z. Fu, X. Cao, J. Wang and X. Sun, "Secure Storage of Data in Cloud Computing," 2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kitakyushu, 2014, pp. 783-786.
- [8]. Amazon Web Services (AWS), Online at <https://aws.amazon.com>, 2008.
- [9]. Microsoft Azure Documentation, online at <https://azure.microsoft.com/en-in/overview/what-is-cloud-computing/>
- [10]. Likhwa MLOTSHWA, Awie LEONARD, Felix NTAWANGA "A Conceptual Framework for Cloud-Computing Management: An End-user Environment Perspective" ISBN: 978-1-905824-51-9
- [11]. H. Hasan and S. Chuprat, "Secured data partitioning in multi cloud environment," 2014 4th World Congress on Information and Communication Technologies (WICT 2014), Bandar Hilir, 2014, pp. 146-151.

BIOGRAPHIES

	<p>Mr. Ajay Kumar Lebaka, pursuing M.Tech (CST) from Baba Institute of Technology and Sciences, Vizag affiliated to JNTUK and completed B.Tech (CSE) from Avanathi Institute of Engineering & Technology, Vizag affiliated to JNTU.</p>
	<p>Mr. Ratna Kumar Jala, B.Tech, M.Tech, (M.Th), (MBA), (Ph.D), is currently working as Associate Professor in Baba Institute of Technology and Sciences, Vizag affiliated to JNTUK and Head of Department of Computer Science Engineering. He has a vast experience of 13 years in this field. Currently he is pursuing his Ph.D.</p>