Data Security in Cloud Computing

[1]Kiran K, [2]Mariadas Ronnie C P

[1]Under Graduate Student, [2]Assistant Professor
Department of Computer Application
SCMS School of Technology and Management, Kochi, India

***Abstract*: Cloud computing is a rapidly evolving model that has many features that are keep on updating. Security of cloud-based applications and data is one amongst the key concerns of cloud customers. In order to provide maximum security, we need a proper software and software management cycle. The knowledge security of cloud systems rests on the classical principles of confidentiality, availability, and integrity, but applied to distributed, virtualized, and dynamic architectures. This paper presents an analysis of data security of cloud environment. Solution exists for some. Analysis of those solutions will determine the lacunae within the data security issues**

***Keywords*: Cloud, types of cloud storage, data security, Cloud services, data security methods**

I. INTRODUCTION

Cloud computing is the process of delivery of many computing services.it includes the delivery of services such as storage, database etc. The term often describes data centers available to several users over the web. Apart from that there exists large cloud with which they share the functions from a central server to many other locations. If the connection to the user is comparatively close, it should be designated as grip server.

With the cloud computing technology, users can use many things that come under the internet and many devices such as laptops and pc's by cloud computing providers.

The main Advantages of the cloud computing technology include cost savings, high availability, and easy scalability.

Cloud storage can be divided into five categories (referred from [1]) in practical applications, namely, public cloud storage, personal cloud storage, private cloud storage and hybrid cloud storage.

1. Public cloud

Public cloud storage is a cloud storage model that enables the user and a particular organisation to store data, manage data and edit data. This type of cloud storage resides on remote cloud server and it is well accessible. The advantages of public cloud such as flexibility, scalability and cost saving attract plenty of small and medium enterprises.[1][4]

2. Personal cloud

A form of cloud storage that stores the individuals' data in the cloud storage and makes the data accessible from anywhere over the internet. Personal cloud storage also helps in the syncing and sharing of the data between multiple devices such as mobile phones, tablets and computers. Personal cloud storage is also known to as mobile cloud storage or pocket cloud storage.[1]

3. Private cloud

Private cloud provides computing services to a private network **(**within the organization**)** and users instead of the general public. Through firewalls and internet hosting private cloud provides high level of security and privacy to the data in the cloud. It also makes sure that both operational and sensitive data are not accessible to third-party users.[1]

4. Hybrid cloud

Hybrid cloud is none other than the combination between private cloud and one or more public cloud, with proprietary software that helps in the communication between each distinct service. A hybrid cloud strategy provides businesses with greater flexibility by transferring the workloads which will be more efficient and cost fluctuate[1]

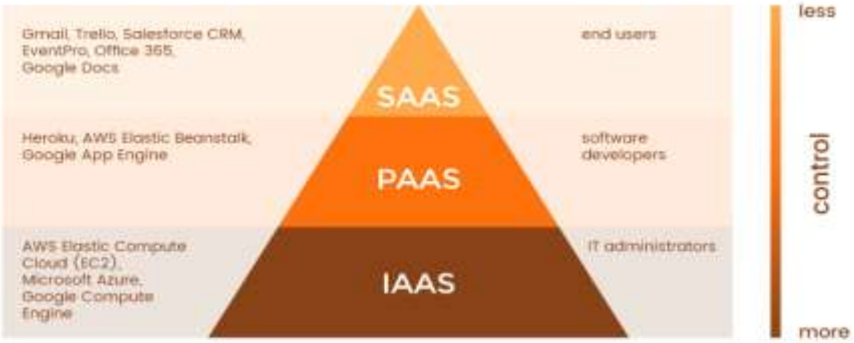
Cloud storage is based on virtualization infrastructure and much similar to cloud computing in the terms of interfaces, scalability and measurement resources. It consists of\four layers they are storage layer, primary management layer, application interface layer and access layer. Apart from this there are certain cloud computing services they are PaaS, SaaS, IaaS [1],[3],[4],[5].

- PaaS (platform as a service): - **PaaS** (Platform as a Service), as the name suggests, provides you computing platforms which typically includes operating system, programming language execution environment, database, web server etc. Platform as a Service (PaaS) provides the runtime scenario. It allows programmers to create, test, run, and deploy the web applications. We can purchase these applications from a cloud service provider on a payment basis and access them using the Internet connection. In PaaS, backend scalability is handled by the cloud service provider, so end- users not need to concern about managing the infrastructure [1],[3],[4],[5].

- SaaS (software as a service): - Software-as-a-Service, or SaaS for short, is a cloud-based method of providing the software to users. SaaS users mainly subscribe an application other than purchasing it and installing it. Users can log into SaaS application from any devices and use it over the internet [1],[3],[4],[5].

- IaaS (infrastructure as a service): -infrastructure as a service (IaaS) is a type of cloud computing service that offers essential compute, storage and networking resources on demand, on a pay-as-you-go basis. IaaS is one among the types of cloud services, along with software as a service (SaaS) and platform as a service (PaaS) . Migrating your organisation's infrastructure to an IaaS solution helps you to reduce the maintenance of on-premises data centres and save money on hardware costs and gain real-time business.

The architecture of cloud computing service is given below


Gmail, Trello, Salesforce CRM, EventPro, Office 365, Google Docs
SAAS
end users
Heroku, AWS Elastic Beanstalk, Google App Engine
PAAS
software developers
AWS Elastic Compute Cloud (EC2), Microsoft Azure, Google Compute Engine
IAAS
IT administrators
less
control
more

II. LITERATURE REVIEW

Jingli Ren on 2020 published a survey paper about data security and privacy in cloud computing. that deals with the cloud security features and the issues that arises.,

Belal Ali, Mark A Gregory, And Shuo Lee on 2019 made a study on multi edge computing architecture that provides the security analysis of the data in cloud platform. Also stated about security technique's and it's about issues of security that occurs n cloud computing.

III. DATA SECURITY

Data security refers to the process of protecting data from unauthorized access and data corruption on a cloud computing platform. Data security includes many methods that provide maximum security and privacy to the data in the cloud.

The various data security methods are [1],[2],[3],[5]

A. IDENTITY AND AUTHENTICATION
B. ACCESS CONTROL
C. DATA ENCRYPTION
D. DATA MASKING
E. DATA INTEGRITY

These are the various methods to provide maximum security to the cloud. Each one of these have different security functions.

A. Identity and authentication

Authentication, along with authorization, is one of the recommended ways to increase data security and protect against data corruption [1]. Authentication technology verifies if a user's details match that are previously entered and helps to approve the authorization. In order to identify the authorized user, we can have passwords, pin etc.…

In other words, it can be said as, Identification is the claiming of an identity. This only needs to happen once on authentication or access process. Any of these three common authentication factors can be subjected for identification. Once identification has been performed, the authentication process must happen. Authentication is the act of verifying or proving the given identity. The issue check both, that such identity actually exists within the known accounts of the secured scenario and also ensuring that the human giving the identity is the correct or valid. The advantages of identify and authentication are

- to enable the user's right to access the system and information
- protect from theft and fraudsters.

B. Access control

Access control is an important component of data security that notices who's allowed to access and use information and resources. Authentication and authorization happen by the process called access control.
There are three types of access control

- Discretionary access control (DAC) is a model of access control that is being determined by the owner of the resource in question. The owner of the resource can decide who have the access and who doesn't have, and exactly what access they are allowed to have. [1][2]

- Role-Based Access Control (RBAC) is a security model in which users are granted access to resources based on their role in a company. RBAC, if implemented correctly, it can be an effective method of enforcing the principle of least privilege.[1][2]
- Mandatory Access Control (MAC) is system-enforced access control mainly based on a particular subject's clearance and an object's labels. Both Subjects and Objects have clearances and labels, respectively, confidential, secret, and top secret.

This determines the type of security that access control can provide

C. Data encryption

Data encryption is a type of data security that is done by using an algorithm (called a cipher) and an encryption key to turn normal text into encrypted ciphertext. To an unauthorized person, the cipher data or text that is present, will not able to read if he/she doesn't have a proper key.[1],[3],[4],[2]

That data can only be decrypted by a user who have the authorized key. Encryption is used to protect the data from security threats

How encryption works?[1]

Encryption is the process of taking plain text, like text message or email, and converting it into an unreadable format called "cipher text." This protects the confidentiality of digital data either stored on computer systems or transmitted through a network on the internet.
When the recipient accesses the message, the information is translated back to its original form. This is called decryption.
To unlock the message, both the sender and the recipient have to use an encryption key —that is a collection of algorithms that convert data back to a readable format.
The two types of encryption systems are:

- Symmetric encryption: - it uses a single password to encrypt and decrypt data.
- Asymmetric encryption: - it uses two keys for encryption and decryption. A public key, which is shared among the users will encrypts the data. A private key, that is not shared will decrypt the data.

D. Data masking

Data masking is the process of protecting sensitive data from thefts and or any unauthorized persons. It is a way to create a fake, but a realistic version of your data.
The main reason for applying masking to a data is to protect the data because the data can be more sensitive and reserved for intended purposes. In order to undertake valid test cycle the data should be remain usable. It must look real and consistent. Although it is a very common factor that the data can be masked
There are two types of data masking:

- Static data masking (SDM) permanently removes the sensitive data by adjusting the data when it is at rest within the database copies that is being provisioned to DevOps environments

- Dynamic data masking (DDM) temporarily hides or replace sensitive **data** in transit and leaving the original at-rest **data** that is unaltered

E. Data integrity

Data integrity means preserving information integrity in a cloud system [1],[2],[3]. The data should not be alter or modified by the unauthorized users. Data integrity is mainly used to provide cloud computing service such as SaaS, PaaS, and IaaS. ... Cloud computing providers are requested to maintain data integrity and accuracy.
Integrity of data is important because it assure the protection of searchability and traceability of the data to their original source. Data performance and stability will also increase when you assure effective data accuracy and protection of data. Maintaining the integrity of data and ensuring the completeness of data is very important.
There are two types of data integrity

- Physical integrity: - Physical integrity is meant by the protection of data accuracy as it's stored and retrieved. When natural disasters strike, power goes out, or hackers corrupts the database functions, physical integrity is compromised. Human error, storage erosion, and a host of other issues can also make the data processing impossible for system programmers, applications programmers, and internal auditors to obtain accurate data.

- Logical integrity: - Logical integrity keeps the data unchanged as it's used in various methods in a relational database. Logical integrity protects data from human errors and hackers, but in a much various way other than physical integrity does.

IV. COMPARATIVE STUDY ON METHODS OF DATA SECURITY

There are various methods that provide maximum security to the data present in the cloud. The methods that had given above tells various security methods and their contribution for security of data.
Authentication is the method to boost data security and access control is the reaching of information and resources. By entering the correct details only, the user can enter to the system and it checks whether the information that have entered is correct. If the information given is wrong then they block the login of the user. After that when the information or data that have been given by

us is correct it then undergoes the access control and allow the user tot login to the system, they also prevent the entry of unauthorized user.
After login, the data will be encrypted which means that the authorized user can only view the data that have been given. The user will have a key which is a cipher text so only by using the key one can access it. if anyone who doesn't have the key can't go through the data. Even though anyone who enters with the username and password can't go through the data if he/she do not have the key.

Data masking when compared to the encryption it is more secure because the data that has been masked and the results can be permanent (no need to reverse the masking). Data masking is a very fine level security approach to protecting data attributes.
In data masking the original data can be masked. If an unauthorized user enters the system, he/she can't see the data because the data is masked by another name which is real. It is very hard to hack a data that is masked. In the case of data masking if any hack appears it will be very difficult to understand the real and fake data, but for encryption it is easy to have the key and can hack so if any hack occurs the encrypted data can be decrypted.

Data integrity is important because it ensures the searchability and traceability of the data to its original source. Data performance and stability also increase when you give effective data accuracy and data protection. Maintaining the integrity of data is essential. It is important because when something happens to the data or if it is lost the data will be present in their original source so if a hack or any unauthorized malfunctions happen the data would be there itself. It also provides the accuracy of data. When compared to data masking the integrity of data is to maintain the data in its original source when a data is deleted or hacked. so, at that time the hacker can able to see the information in it, but data masking as it is a irreversible process one cannot steal any information.

V. THE BEST METHOD

Theoretically speaking, Data masking is the best method among them because it provides maximum security to the data. If any unauthorized user enters to the system there will be a couple of data which will be masked that means an organizational data is faked but it looks the same. So, when any external user enters to the system, he/she can't be able to recognize true data. In the case of encryption, the data cis vulnerable and even though it is hacked one can decrypt the code and access the data but when it is masked if any hack occurs one cannot identify which is original and masked and can't easily access the data. Even though data masking is irreversible it protects the data in order to achieve the anonymization of the data
It has the more ability to hide sensitive data.

VI. CONCLUSION

In this paper we discussed about the data security in cloud computing. As we know that it gives how significant the data security is. We should provide more security in cloud because highly sensitive data is available there. So, this paper deals with various methods that offers the privacy of the data. Data security is an important factor in a cloud server. There are many threats for data privacy and confidentiality like data leakage, hacking, misuse etc. so data security is very essential in cloud system and modern IT world

V. ACKNOWLEDGEMENT

I would like to acknowledge and thank my Guide Mr Ronnie D Mariadas for his support and Guidance in helping me complete this research seminar.

REFERENCES

[1]. P. Yang, N. Xiong and J. Ren, "Data Security and Privacy Protection for Cloud Storage: A Survey," in IEEE Access, vol. 8, pp. 131723-131740, 2020,: 10.1109/ACCESS.2020.3009876.
[2]. BELAL ALI , (Member, IEEE), MARK A. GREGORY , (Senior Member, IEEE), AND SHUO LI , (Member, IEEE) School of Engineering, RMIT University, Melbourne, VIC 3000, Australia Corresponding author: Mark A. Gregory (mark.gregory@rmit.edu.au)
[3]. P.A, Adeeb. (2014). A Seminar Report on SECURITY IN CLOUD COMPUTING.
[4]. Fakhruddin Noori , Abdul Ghafar "Omerkhel", 2021, A Review on Data Security in Cloud Computing, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 03 (March 2021),
[5]. E. Poonguzhali, Suhas Rao M V, Shanth Gk, Mujasem Khanum, 2017, Protection and Security of Data in Cloud Computing, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICPCN – 2017 (Volume 5 – Issue 19),