ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

Review on Data Security and Encryption in Cloud Computing

Amar Kumar Choudhary[1] and Prof. Dr. Anand Mohan[2]

[1] Research Scholar, MIT College of Management, MIT Art, Design and Technology University, Pune.

[2]Professor, MIT College of Management, MIT Art, Design and Technology University, Pune

Abstract

Information and data governance relies heavily on data security as an instrumental means of enforcement. Its application, like all aspects of cloud computing security, should always be uncertain, as it is not feasible to prevent all that uniformly. Whether or not the cloud is involved, this is true for data security in general. However, many companies are not used to sharing such sensitive information with outside parties or to combining all of their private information into a single pool of resources. As a result, rather than adopting a risk-based strategy, which is safer and more practical, the natural reaction would have been to develop a general security measures for "anything on the cloud." Encrypting everything in Saas, for instance, can be a sign that an organisation should not employ a certain vendor in the beginning if it doesn't trust it. Encrypting all data, however, is not a foolproof solution and might give one a misplaced sense of assurance, as in the case of encrypting data flow without additionally checking the security of the hardware. Data security and information security are commonly conflated, but for the purposes of this study, we shall concentrate on the regulations pertaining to the data's protection, with encryption as one of the most important.

Key Words: Cloud Computing, Security Issue, Cloud Challenges, Data Security, Encryption

1.1 Overview

Data security refers to the measures taken by an organisation to safeguard the personal information it holds and guard it from security incidents and data breaches.

Cloud encryption is the process of altering or encrypting data before it is transferred to cloud storage. Unencrypted data (such as a text, file, code, or image) is transformed into an unintelligible form (cypher text) so that it can be concealed from unauthorised and malicious users. This is the simplest and most crucial method to prevent cloud data from being stolen, hacked, or read by someone with malicious intent. Companies that provide cloud storage encrypt data and provide encryption keys to consumers. These keys are used to securely decrypt data as needed. The process of transforming encrypted data into readable data is known as decryption.

1.1.1 Data Security Controls

There are three types of data security controls. In this section following points are covered:

1. Managing which data is uploaded to the cloud (and where).
2. Managing and protecting data in the cloud. The following are the key controls and processes:

a) Access controls
b) Encryption
c) Architecture
d) Alerting/monitoring (of usage, configuration, lifecycle state, etc.)
e) Additional controls, such as those associated with the product/platform/ service of the cloud provider, data loss protection, and organization access control.

3. Lifecycle management information security

a) Data residency and location management.
b) Ensure compliance, which includes audit artefacts (logs, configurations).
c) Business continuity and backups

1.1.2 Cloud Data Storage Types

Compared to conventional storage solutions, cloud storage may support a greater variety of data storage formats because it is virtualized. Although these may use conventional data storage methods behind the virtualization layer, the actual virtualization technologies used to provide users with access to cloud storage are likely to be new. The most typical are the following:

a) Object storage: A file system functions similarly to object storage in terms of operation. Typically, "objects" are files that are then stored in a manner specific to each cloud platform. Although cloud service providers might very well provides front-end interfaces that support standard file sharing protocols, the majority of access is made possible using APIs.

b) Volume storage: For instances/virtual machines, volume storage is effectively a virtual hard disc.

c) Database: Cloud service provider and platforms may offer a number of databases, including their own proprietary systems, as well as commercial and open-source alternatives. Proprietary databases typically have separate APIs. The service provider hosts open-source or for-profit databases that frequently use established standards for connections. Databases can be either file-based or relational, with NoSQL and other key-value stores falling under the latter category (e.g. HDFS).

d) Application/platform: Application/platforms include a (CDN) content delivery network, documents saved in SaaS, caching, and other unique options.

In addition, the vast majority of cloud systems employ redundant, long-term storage mechanisms that periodically employ data distribution (Data fragmentation or bit splitting are other terms for the same thing). This method splits up huge data chunks into many copies on various physical storage devices and ensures excellent durability. The data is consequently physically dispersed.

1.1.3 Cloud Data Migration Management

Prior they can safeguard their data in the cloud, the majority of organisations want a solution to manage data that is kept in private or public cloud providers. This frequently has just as much bearing on compliance as it does on security.

Establish rules governing the sorts of data that are allowed and the locations in which they are permitted, and these rules are then linked to the fundamental security standards. Examples of such policies include "Personally Identifiable Information (PII) is allowed on x services provided that y encryption and access control restrictions are met."

Next, find the locations of the organization's crucial data repositories. Utilising techniques such as Database Activity Monitoring and File Activity Monitoring, keep an eye out for significant migrations and activity. This

is essentially how the "early warning system" for large data transfers is made, and it also serves as a vital data security check to find serious security breaches and abuse scenarios.

Watch any data transfers and cloud consumption to find actual migrations. In order to do this, an organisation can use the following resources:

CASB: Cloud Access and Security Brokers, sometimes known as Cloud Security Gateways, use a number of techniques to identify internal cloud service usage, including network monitoring, integration with third-party network gateways, and even DNS query monitoring. Most of these systems first identify the offerings users are connecting to, and then they provide monitoring of activity on allowed services through API calls (where available) or inline interception which is called man in the middle monitoring. DLP and other security alarms are enabled by many cloud services (SaaS, PaaS, and laaS), and some even provide controls to better regulate the usage of sensitive data in cloud services (SaaS, PaaS, and laaS)[6].

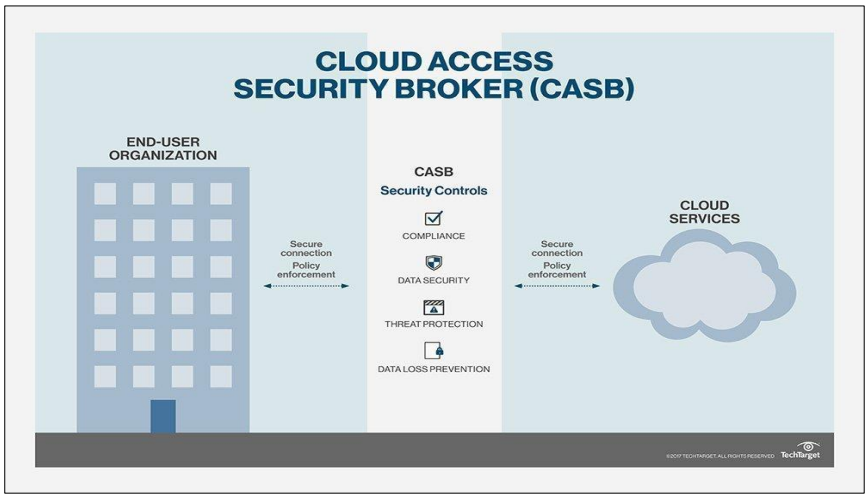

Figure 1: Cloud Access Security Broker (CASB) [1]

URL Filtering: A URL filter or web gateway, while less effective than CASB, can assist an organisation in determining which cloud technology individuals are now using or attempting to use.

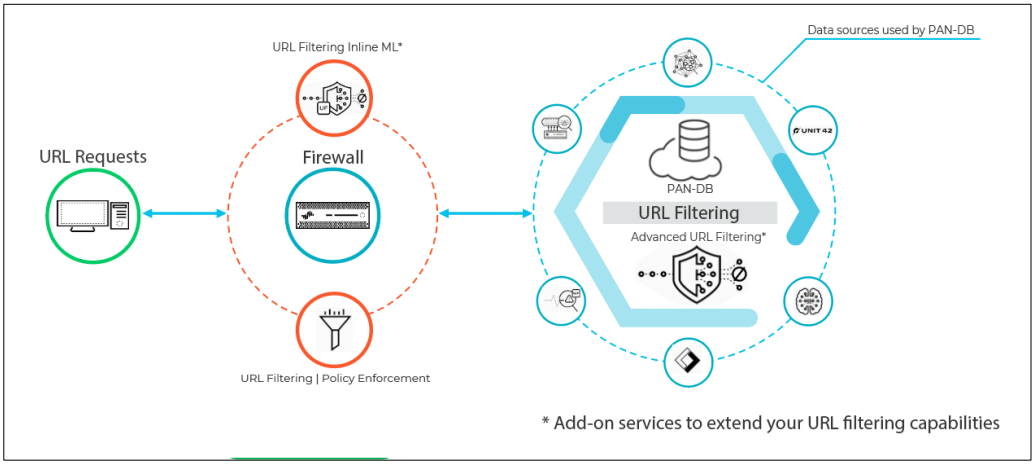

Figure 2: How URL filtering works - URL Filtering Capability [2]

DLP: If a company monitors online traffic, a Data Loss Prevention (DLP) tool can assist in identifying data transfers to cloud services (and peeking within SSL connections). On the other hand, certain cloud SDKS and APIs may protect sensitive data and traffic in a manner that DLP technologies cannot read, preventing them from understanding the payload. DAM facilitates data migrations flow to the cloud (Discovery). [6]

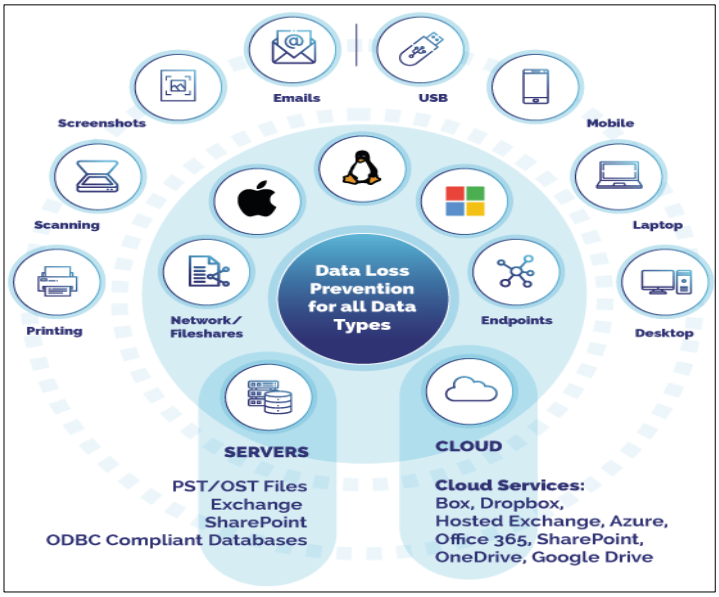

Figure 3: Data Loss Prevention Monitoring [3]

1.1.3.1 Securing Cloud Data Transfers

Make sure the data is secure while it is sent to the cloud. Since using provider systems is often more cost-effective and safe than "manual" data transfer procedures like Secure File Transfer Protocol (SFTP), it is crucial to understand the provider's data migration mechanisms (SFTP). An example: Setting up a personal (SFTP) Secure File Transfer Protocol server on a virtual environment hosted by the same service provider is almost certainly less stable and safe than sending data to the object storage of the provider via an API.

Several different modes of in-transit encryption are available, each tailored to a specific cloud service. Another option is to use encryption on data before uploading it to the cloud (client-side encryption). Encrypting data transfers over a network (using TLS/SFTP, etc.) is yet another possibility.Transport Layer Security (TLS), which is a crucial security element, is typically included as standard in cloud provider APIs; if it isn't, choose a different provider. A third choice would be proxy-based encryption, which includes putting an encryption proxy in a trusted geographical area between the cloud computing user and the cloud service provider in order to maintain encryption prior to data transmission.

Organizations may occasionally be compelled to accept data that is unreliable or available to the public. Make sure there are security procedures in place to purge data before it is processed or integrated with existing data if it accepts data from partners or the general public. Before combining this information, it is imperative to first isolate it and scan it.

1.1.4 Securing Data in the Cloud

The essential data security controls across all technologies are access controls and encryption.

1.1.4.1 Cloud Data Access Controls

A minimum of three layers of access controls should be implemented:

- **Management plane:** These controls are used to control user access to the management plane of the cloud platform. For instance, a user can access data kept in object storage by logging into the web portal of a laas service. Thankfully, the majority of cloud systems and providers include default deny access control settings.
- **External and internal sharing controls:** An additional layer of restrictions will be implemented if data is shared externally with the public or partners who do not have direct access to the cloud platform.
- **Controls at the application level:** On the cloud platform, you can build custom applications as well as design and implement access controls.

Depending on the cloud service model and provider-specific features, access control methods will vary. Make a matrix of entitlements based on the capabilities of the platform. The resources and features that should be made available to particular users, groups, and roles are laid out in an entitlement matrix.

Entitlement	Super-Admin	Service-Admin	Storage-Admin	Dev	Security-Audit	Security-Admin
Volume Describe	X	X		X	X	X
Object Describe	X		X	X	X	X
Volume Modify	X	X		X		X
Read Logs	X				X	X

Figure 4: Entitlement Mappings and Fine-Grained Access Controls [6]

Validate control frequently (ideally continuously), paying close attention to any publicly traded shares. Set up notifications for any updates to permissions that permit public access, new public shares, or both.

Depending on the technology, the projected entitlements' scope will vary greatly. While some databases offer row-level protection, others only permit open access. Whereas others will solely rely on the cloud storage platform on its own, which will run in virtual machines, others will enable link entitlements to the cloud platform's identification and enforcement mechanisms.

Understanding potential outcomes, outlining them, and developing a matrix are essential. Of course, this covers more than just file access; it also covers databases and all cloud data storage systems used by the organisation.

1.1.4.2 Storage (At-Rest) Tokenization and encryption

Encryption options are influenced by service models, providers, and application/deployment specifications. Since key management is just as crucial as encryption, it will be covered in the next section.

Tokenization is not the same as encryption, as both are, technically speaking, two different things. Data is protected by encryption by utilising a mathematical method to "scramble" the information, which can only be unlocked (decrypted) with the appropriate key. A text blob in cryptography is the end outcome. In contrast, tokenization replaces the data with a random integer. A secure database is then used to store both the original and the randomised copies for later retrieval.

Tokenization is frequently used when the original data format is critical. For example, replacing new credit card details with the very same formatted string of text in an old system. Maintaining the format Encryption uses a key to encrypt data while keeping the same fundamental structure as tokenization, but, due to compromises, it might not be as cryptographically secure as tokenization.

Data, the encryption engine, and key management make up an encryption system. Of course, the information it encrypts is the data. The engine is in charge of mathematically encrypting data. The encryption keys are finally in the control of the key manager. The location of each of these parts is the main focus of the overall system design. Organizations should begin the process of creating an encryption system with a threat model. Does an organisation, for instance, trust a cloud provider to keep track of its keys? What are the chances of finding the keys?

1. laas Encryption

Depending on data, multiple ways can be used to encrypt laas volumes.

a) Volume storage encryption

i) **Encryption handled by the instance:** The encryption engine runs inside the instance, and also the key is kept in the volume but secured by a passphrase or key pair.

ii) **Encryption managed externally**: Although the encryption engine runs in the instance, keys are managed and issued to the instance on demand.

b) **File and object storage**

i) **Client-side encryption:** When object Server storage is used as the back-end for an application such as mobile HSM, SECaaS, VM, or Server applications, encrypt data using an encryption engine built into the application or client.

ii) **Server-side encryption:** After being transmitted, information is secure on the server (cloud). The encryption engine is controlled by the cloud service provider who has access to a secret key.

iii) **Proxy encryption**: The volume is linked to a specific appliance or piece of software in this paradigm, and that instance is then linked to the encrypted data instance. All cryptographic operations are managed by the proxy, which can store keys both internally and externally.

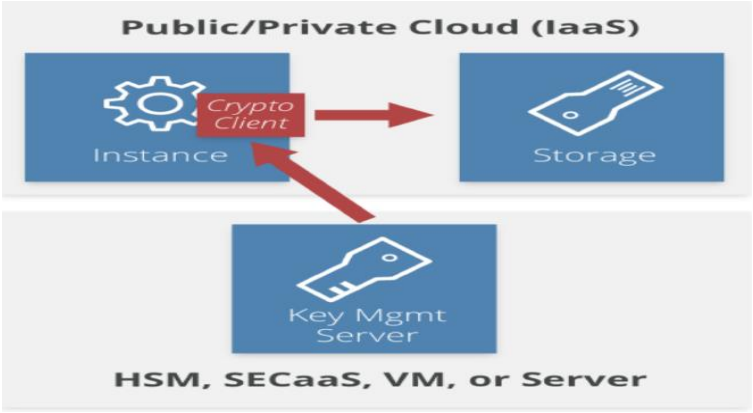

Figure 5: Externally Managed Volume Encryption [6]

2. PaaS Encryption

PaaS encryption varies a lot due to all of the different PaaS systems.

a) **Application Layer Encryption**: Data is encrypted at the application layer in the PaaS application or by the client access platform.
b) **Database encryption:** Using built-in encryption made possible by a database platform like (TDE) Transparent Database Encryption, data is encrypted in the database or at the field level (TDE).
c) **Other:** These are application layers that are maintained by the provider, such as the messaging queue. When laas is utilised for underlying storage, there are also choices.

3. SaaS Encryption

Any of the above listed possibilities are available to SaaS providers per customer keys should be used whenever possible to better guarantee multitenancy isolation. Consumers using SaaS have the following options:

a) **Provider-controlled encryption:** In a SaaS application, data is encrypted and managed by the provider.
b) **Proxy encryption:** Before being delivered to the SaaS service, data passes via an encryption proxy.

1.1.4.3 Key Management (Consisting of Customer-Managed Keys)

The most crucial aspects to take into account when handling keys are performance, accessibility, latency, and security. Can it accomplish this while upholding security and compliance and getting the appropriate key to being in the proper location at the proper time?

There are four possible approaches to key management:

a) **HSM/appliance:** Deliver keys to the cloud using an appliance-based key manager or a standard hardware security module (HSM), which will probably need to be installed on-premises.

b) **Virtual Application/Software:** Integrate a key management system, either as a virtual appliance or as a piece of software, into your cloud infrastructure.
c) **Cloud provider service:** TA crucial management service. Before selecting this choice, be sure to comprehend the security model and SLAs to ascertain whether the key will be disclosed.
d) **Hybrid:** An Organisation can utilise a hybrid approach, such as employing an HSM as the root of trust for keys while distributing application-specific keys to a virtual appliance in the cloud that solely handles keys for that context.

Customer-Managed Encryption Keys

While the provider runs the encryption engine with a client-managed key, a cloud customer can keep track of their own encryption key. For instance, many data is encrypted by default by service providers using keys that they completely control within the SaaS platform. Some people might let you use your own key instead of the encryption system's key. Verify that the vendor's methods adhere to the requirements of the organisation. [7]

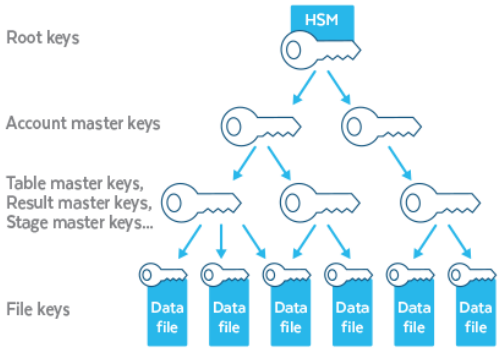

Figure 6: Data Encryption Hierarchy with Customer Managed key [9]

Some providers may need to manage the key using a service provided by the provider. Although the key is kept by the client, it is nevertheless theoretically available to the provider. This does not imply that it is insecure as data storage and key management systems are separate, data compromise would need agreement on the part of numerous personnel at the provider. However, depending on local regulations, a government request could still divulge keys and data. Organisation might be able to keep the keys off-site and only give them along on a request-by-request basis.

1.1.5 Data Security Architectures

Data security is influenced by application architecture. Cloud provider's features may help to decrease the attack surface, but make sure organisation demand adequate infrastructure security. Gap networks, for example, by using cloud storage or a queuing service that interacts through the provider's network rather than its own. As network attack pathways are restricted, attackers must either principally invade the cloud provider or restrict their attacks to application-level vulnerabilities.

Using object storage instead of SFTP for data transfers and batch processing to static instances is an example. Another example is messaging queue gaps, which involves running application components on multiple virtual cloud networks and only connecting them via the cloud service provider's messaging queue service. This prevents network assaults from spreading from one part of the programme to another.

1.1.6 Auditing, Monitoring and Alerting

These ought to be integrated into the broader cloud monitoring system. Identify (and notify about) any changes to sensitive data's public access or entitlements. When tagging is available, use it to assist alerting.

Organisation need to keep track of both storage access and API as data can be accessed via either—in other words, using an API call to access data in object storage or using a public sharing URL to access data in object storage. For example Database activity monitoring could be a viable solution. Make sure logs are kept in a safe place, such as a dedicated account for logging in.

1.1.7 Additional Security Measures for Data

1.1.7.1 Provider-Specific/Cloud Platform Controls

There may be data security controls in place on a cloud platform or provider that aren't covered elsewhere in this area. Although encryption and access control are usually used, these guidance cannot cover all available methods.

1.1.7.2 Data Loss Prevention

(DLP) Data loss prevention is a technique for keeping an eye on and safeguarding the information that employees access by keeping an eye on web, local systems, email, and other traffic. It is more suited to SaaS than PaaS or laaS, where it is less common, and is used less commonly in data centres.

a) **CASB:** Some CASBs have rudimentary DLP capabilities, as well as a rule that states that a credit card information should not be saved in a cloud for the sanctioned services they protect. The tool, the cloud platform, and the manner in which the CASB is incorporated for monitoring all have a substantial effect on efficiency. A few CASB solutions can also send traffic to specific DLP platforms for additional analysis beyond what the CASB provides.

b) **Cloud provider feature**: Providers of cloud storage and collaboration services may offer DLP features, such as scanning uploaded files for potentially malicious data and taking appropriate security precautions.

1.1.7.3 Organizational Rights Management

Similarly to DLP, this is a staff security mechanism that is not always required in the cloud. Existing tools may jeopardise cloud capabilities, particularly in SaaS applications, because all (DRM) Digital Rights Management and (ERM) Enterprise Rights Management systems rely on encryption.

a) **Full DRM:** Comprehensive digital rights management employing a conventionally recognised tool For instance, prior to putting a file in the cloud, permissions can be applied to it. Without some sort of integration, it can interfere with cloud provider features like browser preview or collaboration (which, as of this writing, is exceedingly uncommon).

b) Provider-based control: Utilizing native capabilities, the cloud platform may be able to implement regulations that are very similar to full DRM. For example, the device/user/view versus edit policy allows

some users to read a file in a web browser while others can modify and/or download the content. Instead of merely people, some platforms let you link these limits to specific devices.

1.1.7.4 Data Masking and The Creation of Test Data

All these techniques limit real-time data access in apps and secure data in testing and development environments.

a) **Test data generation**: Generating test data entails building a database from scratch using non-sensitive test information from a "real" database. It can make a data collection that is comparable in size and structure to the source but excludes sensitive information by using scrambling and other randomization techniques.

b) **Dynamic masking:** Data is rewritten dynamically using dynamic masking to hide all or part of the data provided to a user, generally utilising a proxy technique. When presenting a credit card number to a user, it is typically used to secure some sensitive data in apps, such as blocking out all but the last digits.

1.1.8 Enforcing Security for Lifecycle Management

a) **Managing data residency/location:** Disable superfluous locations at times. At the container or object level, use encryption to control access. The data is then safeguarded even if it moves to an unauthorised place until the key moves with it.

b) **Ensuring compliance:** Maintaining compliance requires more than just implementing controls; it also necessitates documenting and testing those procedures. These are "compliance artefacts," which include any audit artefacts it may have.

c) **Backups and business continuity:** When think about RTO (Recovery Time Objective) and RPO (Recovery Point Objective), organisation talk about business continuity. The Recovery Time Objective (RTO) is the amount of time it takes to restore a business after a disruption in order to avoid unacceptably negative repercussions. The Recovery Point Objective (RPO) is the maximum time that data can be lost in the event of a disaster.

1.2 Conclusion

Become Learn about the capabilities of the cloud platform being used by the business. Don't discount the security of the data held by cloud services. It is frequently more cost-effective and safer than making it alone in many situations. It establishes access restrictions and builds an entitlement matrix. Cloud service providers' capacity to enforce will differ.

To monitor data streaming into SaaS, use CASB. For modest PaaS and laaS, it might still be helpful, but for significant migrations, rely on current regulations and data repository security. Based on the threat model, choose the appropriate encryption for your data, business, and technology needs. Use storage and encryption methods that are overseen by the provider. Always prefer to utilise a key controlled by the consumer if available. Discuss Utilizing architecture can help to increase data security. Ensure that monitoring of APIs and data levels is in place, and that logs adhere to life cycle policy and compliance specifications. Guidelines have been established to facilitate the development of reliable security systems and the effective implementation of encryption and key management procedures. Specifications used include ANSI X9.69 and X9.73 as well as NIST SP-800-57.

Acknowledgements: This Paper is about data security and encryption in the cloud computing, while many people and organizations have contributed to the study, I wish to express my gratitude to each one of them for supporting me for this paper. Firstly, I thank the anonymous participants of this study and the acquainted participants who so open handedly permitted me into their businesses who shared their realities with me.

I am ever grateful to my supervisor, Dr, Anand Mohan, for guiding me patiently and efficiently, throughout my doctoral journey. I would like to express my gratitude to my institution, MIT ADT University, and the members of the institute. I am particularly grateful for my panel led by Dr. Chabbi Chavan and finally a big thanks to Dr. Anand Mohan, without who, this research would not have been possible.

References

[1]. https://www.techtarget.com/searchcloudcomputing/definition/cloud-access-security-broker-CASB

[2]. https://www.sunnyvalley.io/docs/network-security-tutorials/url-filtering.

[3]. https://bulwark.biz/data-loss-prevention-dlp/

[4]. https://cloudsecurityalliance.org/blog/2021/10/21/cloud-compliance-frameworks-what-you-need-to-know/

[5].https://pages.checkpoint.com/dome9-guide-to-public-cloud-security-and-compliance-for-financial-services.html?utm_term=cyber-hub

[6]. https://cloudsecurityalliance.org/artifacts/security-guidance-v4/

[7]. https://cloud.google.com/storage/docs/encryption/customer-managed-keys

[8]. https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#set-default-key

[9]. https://www.snowflake.com/blog/customer-managed-keys/