UNIVERSITY
OF TURKU

Security comparison of ownCloud, Nextcloud, and Seafile in open source cloud storage solutions

UNIVERSITY OF TURKU
Department of Computing, Faculty of Technology
Master of Science in Technology Thesis

Author(s):
Md. Ibrahim
Supervisor(s):
Dr. Ali Farooq
Dr. Antti Hakkala
June 2022

Master of Science in Technology Thesis
Subject: Networked Systems Security
Programme: Master's Degree Programme in Information Security and Cryptography, Networked Systems Security
Author: Md. Ibrahim
Title: Security comparison of ownCloud, Nextcloud, and Seafile in open-source cloud storage solutions
Supervisor(s): Dr. Ali Farooq; Dr. Antti Hakkala
Number of pages: 73 pages
Date: June 2022

Abstract

Cloud storage has become one of the most efficient and economical ways to store data over the web. Although most organizations have adopted cloud storage, there are numerous privacy and security concerns about cloud storage and collaboration. Furthermore, adopting public cloud storage may be costly for many enterprises. An open-source cloud storage solution for cloud file sharing is a possible alternative in this instance. There is limited information on system architecture, security measures, and overall throughput consequences when selecting open-source cloud storage solutions despite widespread awareness. There are no comprehensive comparisons available to evaluate open-source cloud storage solutions (specifically owncloud, nextcloud, and seafile) and analyze the impact of platform selections. This thesis will present the concept of cloud storage, a comprehensive understanding of three popular open-source features, architecture, security features, vulnerabilities, and other angles in detail. The goal of the study is to conduct a comparison of these cloud solutions so that users may better understand the various open-source cloud storage solutions and make more knowledgeable selections. The author has focused on four attributes: features, architecture, security, and vulnerabilities of three cloud storage solutions ("ownCloud," "Nextcloud," and "Seafile") since most of the critical issues fall into one of these classifications. The findings show that, while the three services take slightly different approaches to confidentiality, integrity, and availability, they all achieve the same purpose. As a result of this research, the user will have a better understanding of the factors and will be able to make a more informed decision on cloud storage options.

Keywords: Cloud Storage, Cloud storage security, Features, Architecture, Vulnerability, Open source cloud storage, ownCloud, Nextcloud, Seafile

Table of Contents

List of Figures

List of Tables

Abbreviations

IT = Information Technology

DevOps= “Development” and “Operations”

SAN=storage area networks

FC=Fiber Channel

NFS=Network File System

SMB = The Server Message Block

HTTP= Hypertext Transfer Protocol

VM= virtual machine

SaaS = Software as a Service

PaaS = Platform as a Service

IaaS = Infrastructure as a Service

DR=Disaster Recovery

SCSI= Small Computer System Interface

iSCSI= Internet Small Computer Systems Interface

IP = Internet Protocol

CIA= Confidentiality, Integrity, and Availability

AGPLv3 = GNU General Public License, version 3

VMM = Virtual Machine Manager

AI = Artificial Intelligence

GDPR =General Data Protection Regulation

CCPA = California Consumer Privacy Act of 2018

WebDAV=Web Distributed Authoring and Versioning

API= Application Programming Interface

REST = Representational State Transfer

AD=Active Directory

LDAP= Lightweight Directory Access Protocol

DLP=Data Loss Prevention

db= Database

STUN=Session Traversal Utilities for NAT

TURN =Traversal Using Relay around NAT

PHP =Hypertext Preprocessor

REDIS= Remote Dictionary Server

GFS = Global File System (LINUX file system)

CIFS = Common Internet File System

FTP= File Transfer Protocol

SSL = Secure Socket Layer

2FA = Two-Factor Authentication

MFA=Multi Factor Authentication

AES = Advanced Encryption Standard

TLS=Transport Layer Security (TLS)

SAML = Security Assertion Markup Language

PBKDF2 (Password-Based Key Derivation Function 1 and 2)

CBC = Cipher Block Chaining

URL = Uniform Resource Locator

XSS = Cross-Site Scripting

UUID = Universally Unique Identifier

DLL = Dynamic Link Library

SUSE= Software- und System-Entwicklung (Software and Systems Development)

HSM = Hardware Security Module

P2P = Peer-to-Peer

DDoS =Distributed Denial-of-Service

1 Introduction

Cloud storage is one of the most significant improvements in information technology. In recent years, it has developed from a theoretical approach to a practical requirement for both individuals and businesses. Cloud storage is one of today's fastest-growing IT areas. It is transforming how we live and interact as a society. Cloud storage facilitates collaboration and sharing by storing data remotely [1]. Additionally, it has a substantial influence on enterprises. Users may exchange and view data from anywhere, even if they do not access their local storage systems. Cloud storage enables us to deliver storage at a lower price while enhancing safety and stability. A cloud storage system is a collaborative service approach that utilizes many devices, application domains, and service methods.[1] [2]

An open standard cloud platform provides its source code for users to examine how the service processes data and gain insight into its design. In the presence of numerous open-source cloud storage solutions, current frameworks and technologies are primarily centered on function features or a simple approach. As a consequence, end-users face difficulties in making appropriate decisions. This thesis covers a comparative analysis of three popular open-source cloud storage solutions (ownCloud, Nextcloud, and Seafile) regarding functionality, architecture, security, and vulnerabilities to overcome this challenge.[3]

1.1 Background

Although cloud platforms have opened new storage and synchronization options for enterprises, they also introduce unknown risks. Trust difficulties would inevitably arise in the context of corporations and cloud services. Trust is shown as a complicated component comprised of asset control, data ownership, failure prevention, and other factors. While cloud storage is becoming increasingly popular, especially among individuals, institutions, and business organizations, particular security concerns remain [1][2]. Many commercial cloud storage companies exist, including Dropbox, Google Drive, Amazon, Etc. However, those third-party cloud storage solutions pose significant security and privacy risks, especially for storing sensitive information. Security, privacy, and data ownership must all be handled in the development and maintenance of systems that fulfill performance and resource criteria, as well as properly address data ownership and personal data ownership issues. On the other hand, frequent concerns about those questions lead to discussions about implementing open-source private cloud infrastructure. Open-source cloud storage solutions improve security and privacy, data backup, productivity, and efficiency.[1], [2], [3]

1.2 Problem statement

The massive content generation of files, photographs, movies, and other items from various digital devices is driving up demand for storage mediums and cloud services. A storage device that does not have remote connectivity is less valuable than a cloud storage device. Dropbox, Google Drive, Amazon, and many other commercial cloud storage providers offer cloud services. However, those third-party cloud storage solutions constitute a substantial security and privacy issue when storing sensitive data. As a result, there is a significant risk of security and privacy violations [2]. An open-source cloud storage solution improves performance, reliability, and efficiency in data backup and security. End-users of open source cloud storage solutions have access to a vibrant community of industry experts who are constantly working to improve the products. This collaborative effort improves security. Most open-source storage models have the advantage of being free. Open source cloud storage solutions are also multi-environment compatible and allow for source code modifications, giving them more flexibility. [2] [4]

Even though there are numerous open-source cloud storage solutions for constructing private open-source cloud storage, it is challenging for end-users to decide on an appropriate cloud storage solution. Despite widespread interest, limited information is available on system architecture, security techniques, and overall throughput implications when choosing open-source cloud storage solutions. There are still no comprehensive comparisons to help compare services and analyze the impact of platform choices. This thesis investigates a comparative view of the three robust open-source cloud storage solutions, namely ownCloud, Nextcloud, and Seafile, in terms of their features, architecture, security, and vulnerabilities. This comparative study helps users acquire a better understanding of the factors and be better equipped to make an informed selection about cloud storage solutions.

1.3 Purpose and Objectives

This thesis aims to compare the three popular open-source cloud storage options in terms of features, architecture, security, and vulnerabilities issues so that users can better understand them and make better judgments. The following research objectives could facilitate the achievement of this aim.

Research Objectives:

I. Identify the features of owncloud, nextcloud, and seafile.
II. To understand the architecture of these three open-source cloud storage solutions
III. To explore the security features of those cloud storage solutions
IV. To examine the vulnerabilities of owncloud, nextcloud and seafile

1.4 Methodology

This thesis has conducted a comparative study of ownCloud, Nextcloud, and Seafile in open source cloud storage solutions. Data has been collected from online articles, verified for accuracy, and compiled the data into an informative and comprehensible format for everyone who reads this thesis. The research was concentrated on analyzing data from journal articles, conference papers, and web pages from previous studies. The author was focused on four attributes: features, architecture, security, and vulnerabilities of three cloud storage solutions ("ownCloud," "Nextcloud," and "Seafile") since most of the critical issues fall into one of these classifications. All referencing, style rendering and reference organization were done using the Zotero open-source reference management solution. The literature review methodology was applied to extract existing knowledge and analyze it in this comparative study. The following four stages were followed throughout the thesis [5] [6].

I. Problem formulation (Research questions)
II. Literature search
III. Data assessment
IV. Analysis, Synthesis, and Interpretation

Problem formulation: The first stage was to determine the goal of the literature review. The goal of the literature review was to compare the three most popular open-source cloud storage solutions (owncloud,nextcloud, and seafile) in terms of features, architecture, security, and vulnerabilities. This thesis investigated the following key research questions to draw a comparative view of the three open-source cloud storage solutions: owncloud, nextcloud, and seafile.

Research questions:

1. What are the features of the three cloud storage solutions (ownCloud, Nextcloud, and Seafile)?
2. What is the architecture of the three cloud storage solutions that have been chosen?
3. What are security features included in those cloud solutions?
4. What are the vulnerabilities of those cloud solutions?

Literature search: In this stage, research was carried out by searching for materials related to the research topic or investigating questions. Academic publications were found in journal articles, conference papers, and web pages from previous research. The search was performed entirely in English, with no geographical constraints imposed. The search retrieved academic articles from all around the world as no geographic constraints were set during the query. The following keywords were considered throughout the search: cloud storage, feature, architecture, security, vulnerability, owncloud, nextcloud, and seafile. These keywords were chosen because they provided the most comprehensive responses to the research questions. Search keywords (terms) help to identify the thesis's scope and ensure the most comprehensive coverage of literature across electronic databases. Each cloud storage solution was searched independently using four attributes (features, architecture, security, and vulnerability), i.e., "Features of owncloud," to obtain appropriate and relevant articles. In addition to the primary keywords, the operator "security OR vulnerability OR features OR "security difficulties") AND (seafile OR Owncloud OR Nextcloud)" is used to discover similar articles based on the three popular open-source cloud storage systems to form a comparison view. The prominent search engine Google Scholar was utilized to find relevant articles.

Data assessment: Data was filtered to get the most accurate and relevant articles at this stage. The Google Scholar search results were included in the Zotero library. The following data was gathered, organized, and entered into Zotero from the articles:

I. The article's title
II. The article's year of publication
III. The source of the article and the type of document
IV. A list of authors who contributed

V. The article's abstract
VI. The language in which the article was written
VII. The article's intended audience
VIII. Volume, issue, and number of pages of the article
IX. Keywords that were used in the article
X. A link to the article's URL.

These articles were evaluated to check if they met the inclusion and exclusion requirements. The evaluation was accomplished by examining the publication's metadata, such as the title of the article, search terms, and summary, then using the inclusive and exclusive requirements listed as follows. Articles with titles that did not approach the research questions or had information not relevant to the questions were eliminated. When an article's metadata failed to identify its content, the entire text was scanned. Articles that addressed the study (research) questions were chosen for assessment. The meta-information and content of publications that did not specify the questions were excluded. Relevant publications to the research problems were identified after examining the meta information and formulating inclusion and exclusion conditions. The final articles were assessed for thematic analysis using preferred reporting elements [5]. The following were the requirements for inclusion and exclusion.

Inclusive requirements :

I. Articles focusing on cloud storage and open-source cloud storage solutions such as owncloud, nextcloud, or seafile.
II. Articles that describe four attributes (features, architecture, security, and vulnerability) of owncloud, nextcloud, or seafile
III. The articles are included that explain a comparative study on open source cloud storage solutions including owncloud, nextcloud or seafile
IV. Articles containing scientific evidence on open source cloud storage solutions, owncloud nextcloud or seafile
V. Articles published since 2010 were assessed to use updated information

The exclusive requirements:

I. Articles that are not obtainable through the university's web access

II. The content of the articles is irrelevant to the research questions
III. Articles that are not conducted in English while having an English abstract
IV. Expert-level articles on the topic

Analysis, Synthesis, and Interpretation: Data was compiled, summarized, gathered, structured, and contrasted with evidence retrieved from the included articles in the final stage. The gathered data was presented in a comprehensible manner that significantly contributes to the existing literature. The data were analyzed in a precise, useful, and understandable format for anyone reading this thesis. The final articles were reviewed and synthesized in each paragraph and throughout the thesis. Synthesize was performed by rephrasing the study's primary findings and applying them to the investigation to find solutions to the research questions.

1.5 Structure of thesis

There are six chapters in this thesis. The remaining sections of the thesis are organized as follows. Chapter 2 describes the fundamentals of cloud storage, types of deployment models for cloud storage, the basics of cloud storage architecture, and security issues in cloud storage. Moreover, it also describes the benefits of using open source cloud solutions. Chapter 3 demonstrates the extensive comparative analysis of the three open-source cloud solutions based on their critical issues, especially security aspects, and research is carried out based on the current literature review. Chapter 4 describes the significant findings of the comparative study and assesses how they relate to research questions. Future research ideas are also suggested in this chapter. Chapter 5 is about the conclusion of the thesis, and Chapter 6 provides the references used in the research.

2 Fundamentals of cloud storage

The major underlying fundamental issues of cloud storage strategies have been presented in this chapter. Security issues in the cloud storage system have also been described.

2.1 Overview of Cloud Storage

The phrase "cloud storage" can be defined in different ways. However, the majority of them have the same meaning: securely the usage of computer resources on demand. This technique is used to provide information technology services, in which contents are obtained via the web using web-based applications (Apps). Apps can interface with servers, which store all the data and software packages [7]. Cloud storage is a technique for organizing, storing, and processing data that is available via the internet and is controlled on the cloud. Because the cloud storage idea is predicated on web access, information can be viewed from any device that has Internet access. Since businesses are no longer required to use their infrastructure and instead lease it from service providers, they may discover that cloud storage reduces the cost of data management [7], [8], [9].

Cloud storage is a service that enables data to be safely backed up, managed, and accessed remotely, typically via the internet. The most common applications include data backup, disaster recovery, archiving, and DevOps. When compared to traditional storage area networks (SAN), cloud storage offers several benefits to customers, including cost savings and enhanced convenience features[7]. The fundamental structure of Cloud storage is illustrated in figure 1.

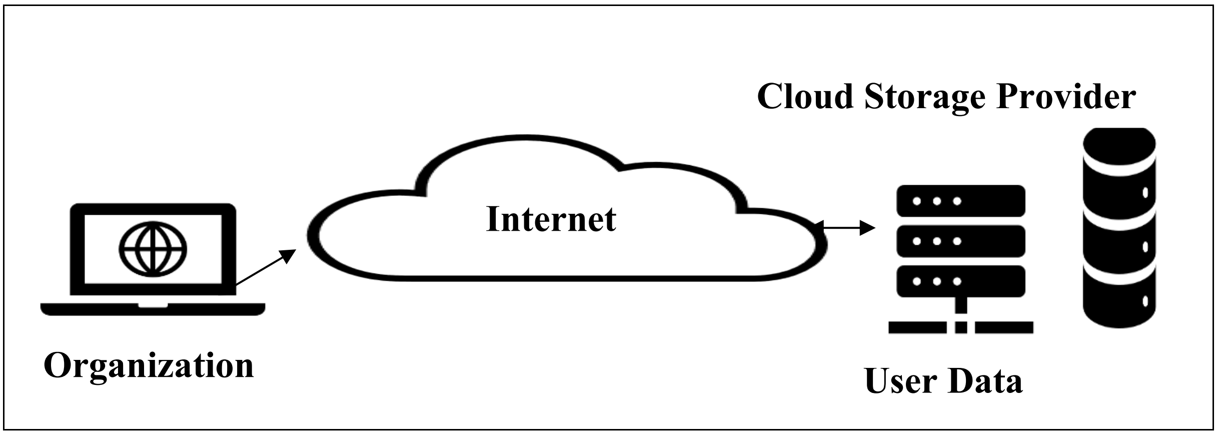

Figure 1: Overview of cloud storage model [9]

Institutions are required to incorporate how they maintain and use their information from formation to end of life caused of massive data volume and the requirements to maintain its security. Cloud storage offers a massive cluster of storage that can be accessed via web services APIs over a stable network connection and instant availability of massive volumes of storage [9]. Cloud storage uses virtualization technology to consider storage devices, servers, and other equipment as a resource pool instead of a discrete system, allowing these resources to be allocated as needed. Integration of application and storage devices is crucial since cloud storage is more than just a memory; it is also a service. The improvement of Cloud Storage from traditional network storage to hosted storage is depicted in Figure 2.

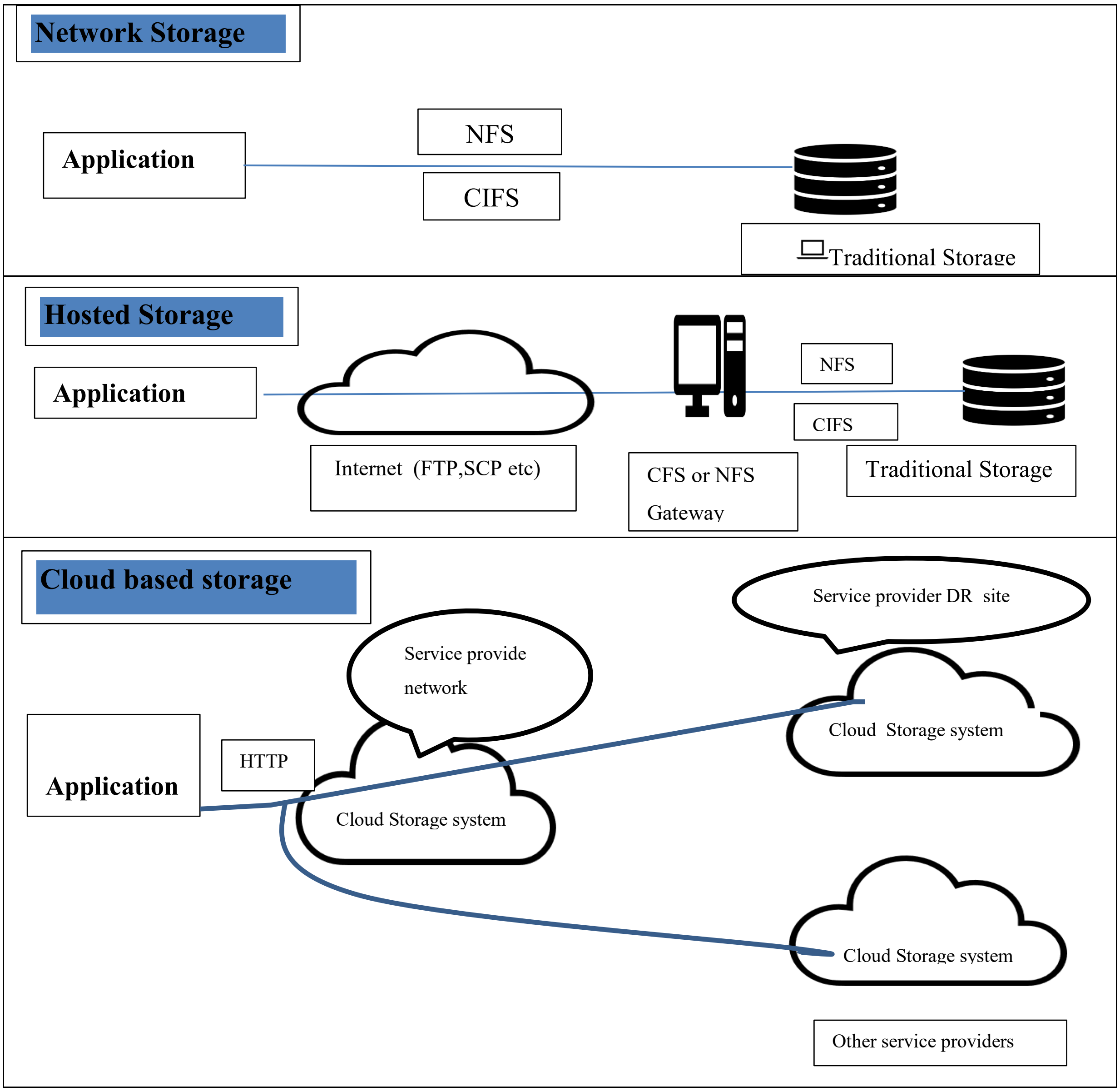

Figure 2: Evolution of Cloud Storage [9]

As a cloud pool of storage resources, cloud storage is used to facilitate various operating modes that can be rapidly implemented in various physical systems and instantly summon resources based on user requirements [8][9]. Cloud storage introduces a new architecture for storing, managing, and analyzing rapidly increasing machine-generated data. Cloud storage will drive new levels of efficiency and economies into enterprise data centers by delivering advanced scalability, management, and the ability to collapse computation and storage on the same processing nodes. Both economic and security advantages come with cloud storage. Physical resources attached to a computer or network are usually more expensive than virtual resources delivered through the cloud. Data is shielded from accidental loss or hardware failures as it is stored on several physical devices. The cloud service continues to operate effectively even if one or more systems collapse since several backups are always preserved on demand. In the event of a system breakdown, data is automatically replicated to other servers [8].

Even though dealing with cloud storage solutions has its drawbacks, one of which is security susceptibility. With the cloud storage option, any confidential information can be exploited through a malicious attack and hackers could take advantage of this information. Instability and other technological issues often make cloud infrastructure insecure. Even the top cloud service provider organizations may experience this kind of issue in terms of maintaining good maintenance standards. Cloud storage is more trustworthy than on-premises computing for a variety of reasons, including high-profile cloud security breaches [2] [10].

2.2 Types of Cloud storage

Cloud Storage consists of two major storage types: Ephemeral and Persistent storage. Ephemeral storage exists when a virtual server is up and running, and ephemeral storage is attached to the virtual server. However, when the virtual server is canceled or down due to errors, the user will lose everything that is running in that ephemeral storage. This type of storage is helpful for scratch, local, and temporary homes for log files. Ephemeral storage is almost always physically attached to the host that a virtual server is running on. Persistence storage persists; it continues to exist, even though a virtual server is not attached to it. The fundamental types of persistence storage are file storage, block storage, and object storage [11] [12]. Figure 3 depicts the differences between each type of storage.

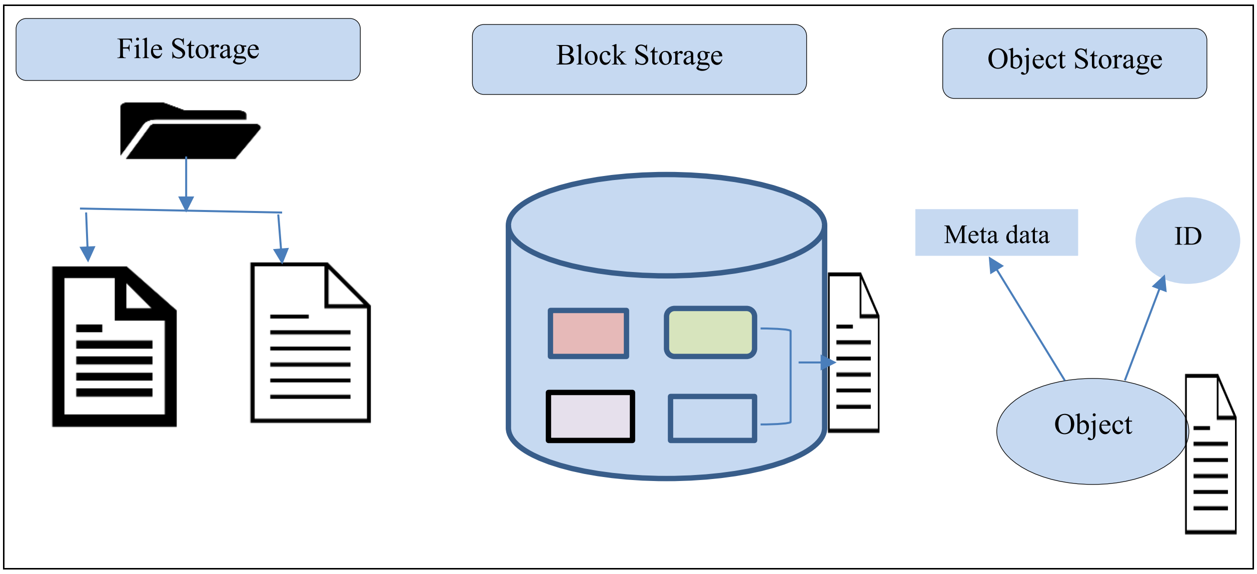

Figure 3: Block, File, Object Storage, and Metadata Layout [12]

The various types of cloud storage are discussed in the following sections.

2.2.1 Block Storage

This category is typically found in a similar data center as the server. It separates data into volumes that are obtained by one or a few servers at the same time. Sectors and tracks represent the configuration of the local hard drive on the volumes[11]. This form of storage is ideal for databases and virtual machines because all communications take place over a dedicated storage area network based on lossless Ethernet or FC technology. It is suitable for all low-latency workloads. One of the disadvantages is the management complexity. Data duplication, compression, and thin provisioning are methods for lowering costs to a manageable level [11] [12]. The architecture of a block storage technique is illustrated in figure 4.

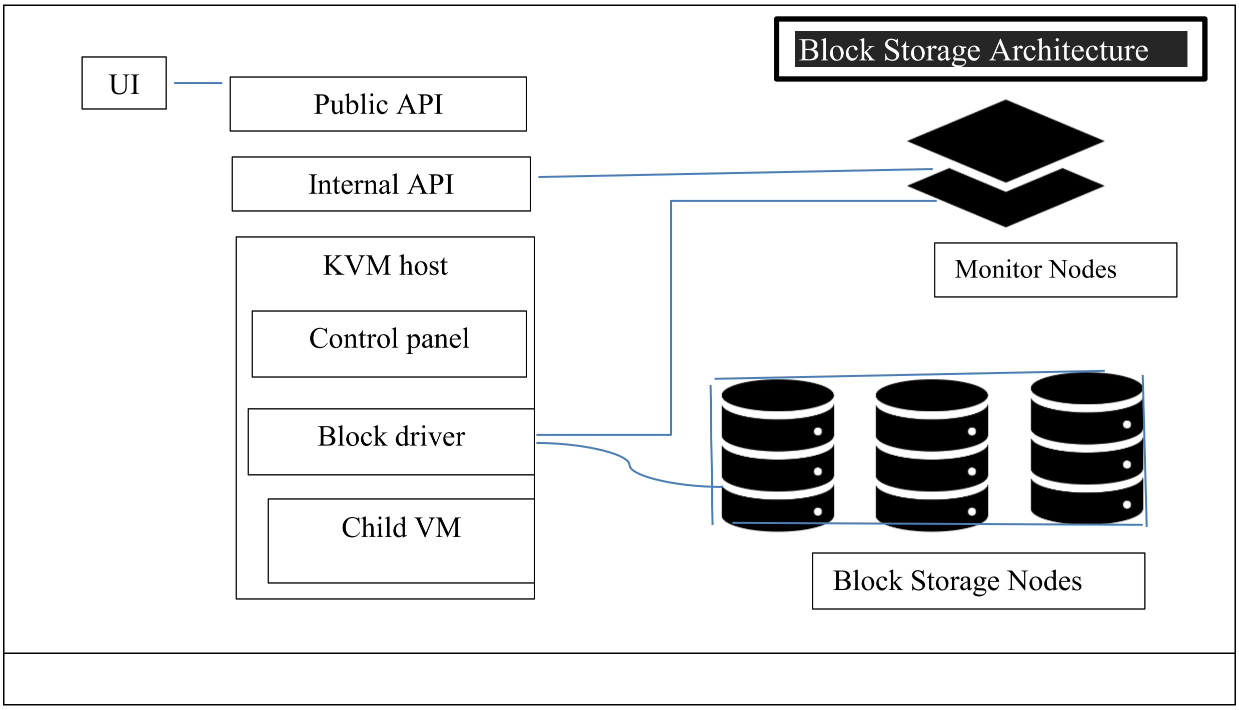

Figure 4: Block Storage Architecture [12]

This technique splits data down into blocks, and every partition is saved by employing a distinctive identifier. It lets users view and edit files. Each block contains a singular address, and its partition is labeled with its address. Metadata is not included in them. Documents are allocated across storage nodes and decomposed into smaller, fixed-size blocks capable of carrying massive quantities of information. This storage type is delivered to computation nodes via high-speed fiber connections, which are generally provided in volumes and placed on compute nodes [11] [12].

2.2.2 File storage

File storage is often given to computing nodes as a Network File System (NFS), which implies the storage is connected to the compute nodes via an ethernet network. This category of storage technique is connected to the local area network. It is accessed by servers as well as other types of clients, such as desktops [12]. This category of storage technique is connected to the local area network. It is accessed by servers as well as other types of clients, such as desktops [11]. Files are stored in hierarchical orders. Files are kept in a single location and are not divided into blocks. Permission is granted to share files and documents with authorization and verification. Additionally, time and date information also includes information on when files were read or generated. The virtual disk storage concept is the most frequently employed in a typical virtualized system [12]. Figure 5 represents an overview of the file storage system.

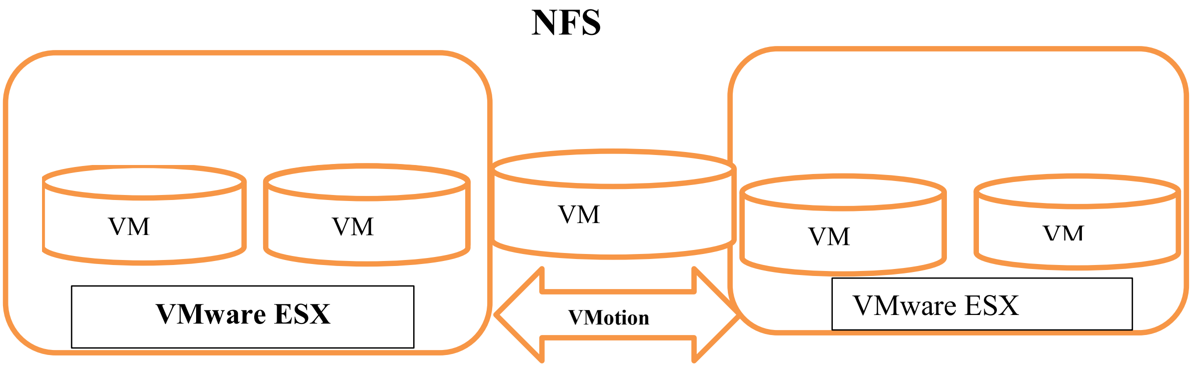

Figure 5: File storage structure [12]

Even though instance or file storage is proscribed in both speed and permanence, it is frequently used for data that has got to be assessed quickly but just for a brief period, like exchanging or paging files. Additionally, it is requested to save information that must be duplicated to many places regularly [11] [12].

2.2.3 Object storage

This type of storage is intended to provide the highest level of accessibility and reliability at scale, allowing any type of device to be connected from anywhere using any type of HTTP-enabled network connection [11]. Data is saved in the form of an object that is made up of data. this type of storage combines the data that makes up a document, includes all of the document's required metadata, and assigns it a unique identity. Figure 6 illustrates the architectural layout of Object storage. Appropriate applications require storage that can be shared across several virtual machines (VMs). Object storage is a technique for splitting data into small, independent segments which are then rehabilitated in a uniform state, with all objects or items at the same layer.

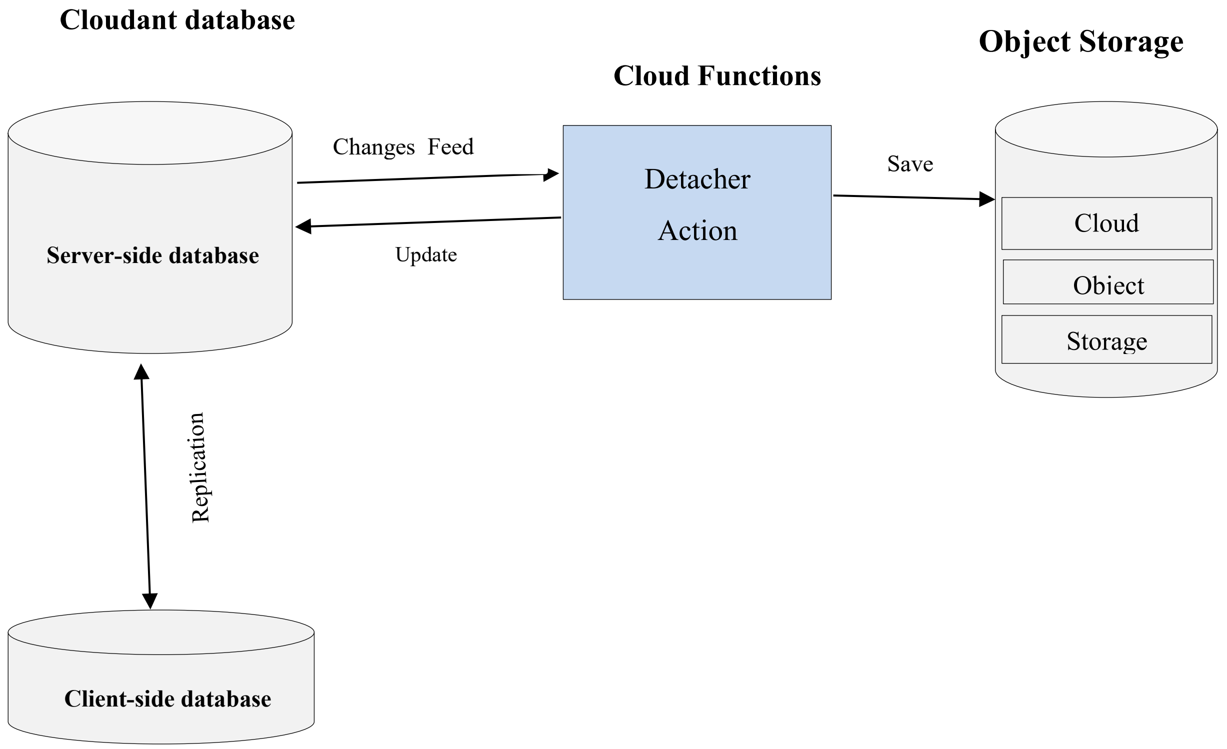

Figure 6: a model of Object storage [12]

In contrast, to file storage, there are no folders or subdirectories. Furthermore, object storage does not consolidate all data into a single document. Each item is identified by a distinctive number rather than a document name and directory. Objects can be archived both locally and remotely on a storage device. Object storage, in contrast to other methods of storage that utilize a file hierarchy structure, stores data as objects. However, long-term consistency is assured. Each object or block is made up of information, metadata, and an identifier unique to it. Object storage is unique in that it seeks to investigate concepts that some other storage systems neglect, such as a directly programmable interface, a namespace, and data propagation [11] [12].

Additionally, unstructured data is preserved in enormous volumes using object storage. This sort of storage is used by music applications, social networking platforms, and online storage services such as Dropbox. One of object storage's several advantages is that it facilitates users to save almost infinite amounts of information. Object storage keeps track of file changes using an HTTP(S)-based interface. Documents are allocated between nodes in this type of storage, which implies that altering a file requires uploading a new version of the full file, which can drastically slow things down. Object Storage is connected through an API and does not demand the use of a compute node [11] [12].

2.3 Service Models of cloud

Each cloud service model supports its unique set of benefits that may be beneficial to a diversity of individuals or businesses. Cloud computing is comprised of three service models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [13]. Figure 7 demonstrates the three cloud service models.

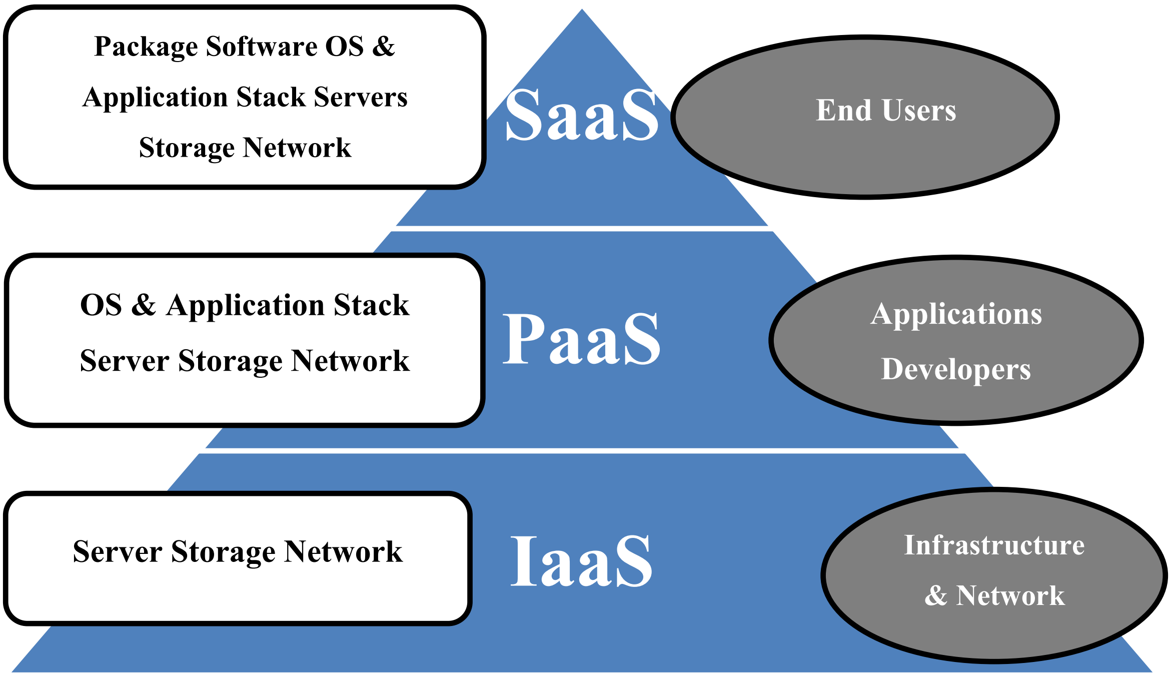

Figure 7: Cloud-computing: The three layers of a generic model [14]

The feature of those deployment models and their advantages are presented as follows.

2.3.1 SaaS (Software as a Service)

SaaS (Software as a Service) approves users to access other publicly accessible clouds of large enterprises for storing their data, such as Gmail. Data security is one of the many benefits that can be derived from a SaaS solution for a company. Using a cloud-based application with centralized data storage can reduce the need for employees to bring sensitive data with them when they travel, which can save duration and money [13] [15]. In the SaaS model, the client is reliant on the serving operator to propose proper security quality. The protection problems associated with SaaS are mostly comparable to those associated with web-based applications. Other security benefits include the sharing of security testing expenses, the enhanced flexibility to deploy secure logs and secure builds, and a system that is more efficiently tailored to perform its functions. It enhances with online tools and apps to perform corporate tasks. End clients are the ones who apply it. It supplies end-users with infrastructure, software as a service, systems, and apps that all exit as SaaS [15] [16].

2.3.2 PaaS (Platform as a Service)

PaaS (Platform as a Service) allows users to host apps or software on other public clouds, such as Google App Engine. Infrastructure, such as networks, servers, storage, and operating systems, are not under the customer's govern in the cloud environment. However, the customer has complete control over the implemented apps and their configurations [13]. It offers virtual platforms and tools for developing, testing, and deploying applications. Developers take benefit of it. It delivers application runtime environments and deployment tools. Infrastructure and Platforms are provided by PaaS [15].

2.3.3 IaaS (Infrastructure as a Service)

IaaS (Infrastructure as a Service) allows customers to virtualize any physical machine and grant use them for a fee. The customer is accountable for preserving control over the protection of the system, which includes the data, apps, and operating software. The developer under the IaaS model has full governance over security, and there are no safety gaps in the virtualization manager [16]. It builds platforms for app development, testing, and deployment and delivers a virtual data center to save information. It's a tool that network architects utilize. Users can access virtual computers, storage, and other facilities through it. IaaS (Infrastructure as a Service) is a service that only delivers infrastructure. [13] [15]

2.4 Types of Deployment models

A brief overview of five cloud deployment models of the cloud is provided as follows.

2.4.1 Public cloud

A public cloud model strategy is used to gain access to excellent accessible and scalable services that are hosted on a distributed infrastructure. These services are available on request, inexpensive, and require no maintenance to operate effectively. Small firms can expand quickly without needing to make huge upfront capital investments or pay high operating costs due to this[17]. The public cloud model is shown in figure 8.

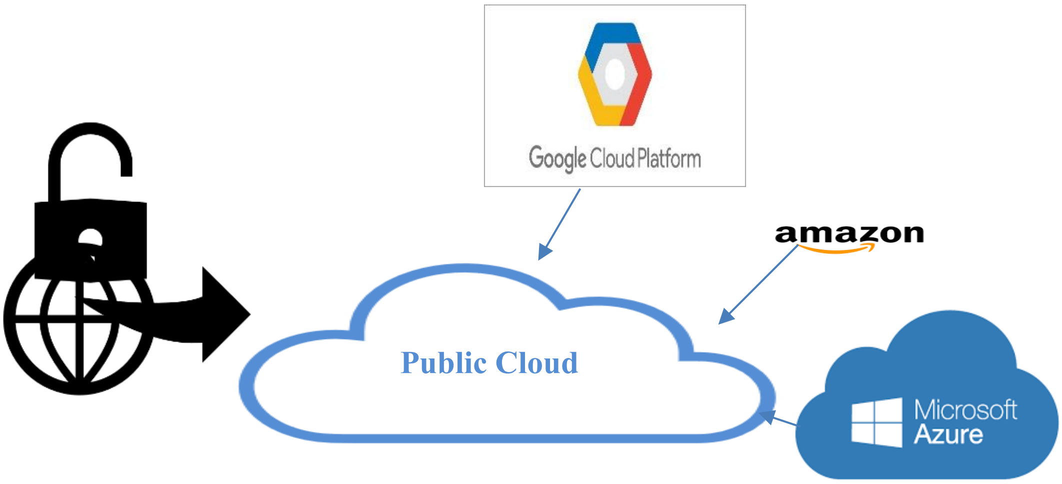

Figure 8: Public cloud Model [17]

The users should ensure the protection and proper configuration of their services to run their applications. privacy and security are significant concerns for many businesses when deploying applications in this deployment model. The users should ensure the protection and proper configuration of their services to run their applications [18] [19].

2.4.2 Private cloud

This deployment model is more common for government agencies, financial institutions, and healthcare institutes to employ this model than for other institutions. A private cloud can be hosted on-premises or through a proprietary service provider [17]. Figure 9 shows the private cloud deployment model.

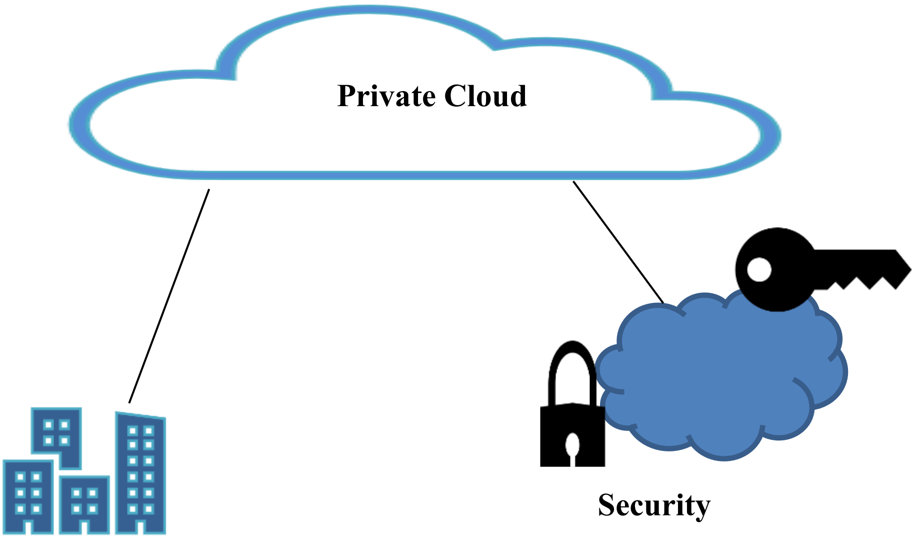

Figure 9: Private cloud model [17]

As for the other components, they are dedicated. Some advantages of using a private cloud include extensive scalability, efficiency, protection, and customizability are just a few of the many benefits of cloud computing. There are possible to get security or performance concerns with private cloud storage [18] [20].

2.4.3 Hybrid cloud

Hybrid cloud storage is a method of cloud storage management that combines on-site and off-site resources. In this model, the services are distributed between the public cloud and the private cloud. The private cloud handles confidential records and critical applications. On the other hand, the Public cloud supports scalable and inexpensive infrastructure, and critical services with no sophisticated information are engaged [17] [18]. The hybrid cloud model is represented in figure 10.

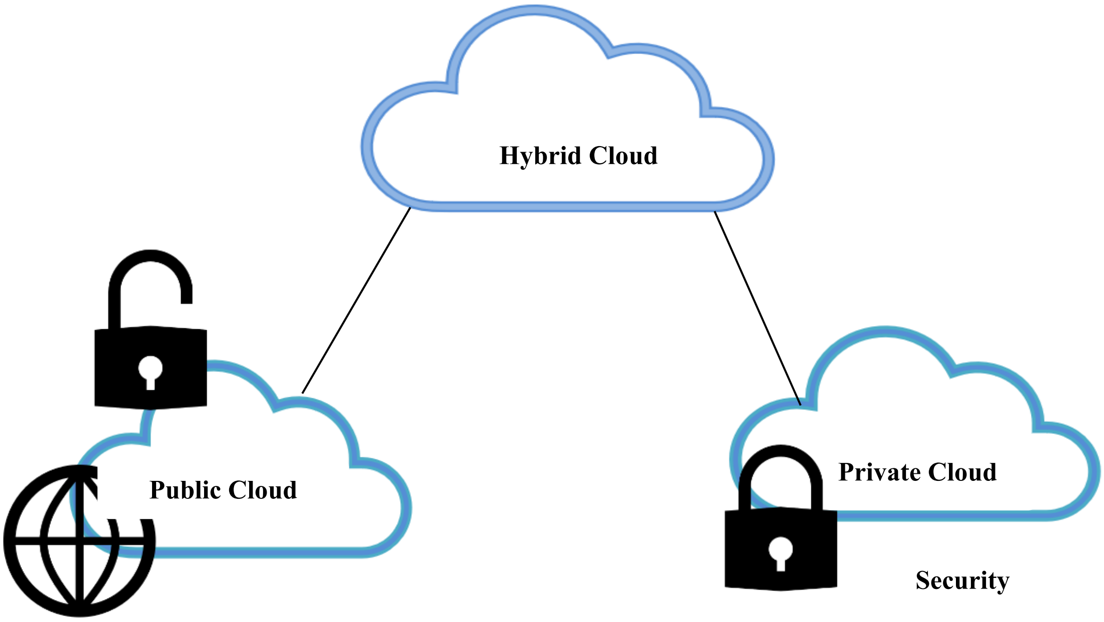

Figure 10: Hybrid Cloud Model [17]

Cloud computing infrastructure hybrid IT has gotten engaged by fusing the benefits of public and private cloud computing. Hybrid cloud storage is a popular option for businesses to simplify data backup and disaster recovery (DR) planning [19] [20].

2.4.4 Community cloud

A community cloud deployment is an improved version of a private cloud architecture that offers a cloud service to certain businesses and organizations. A hybrid cloud that mixes private and public cloud assets is a community cloud [17]. The general community cloud model has been described in figure 11.

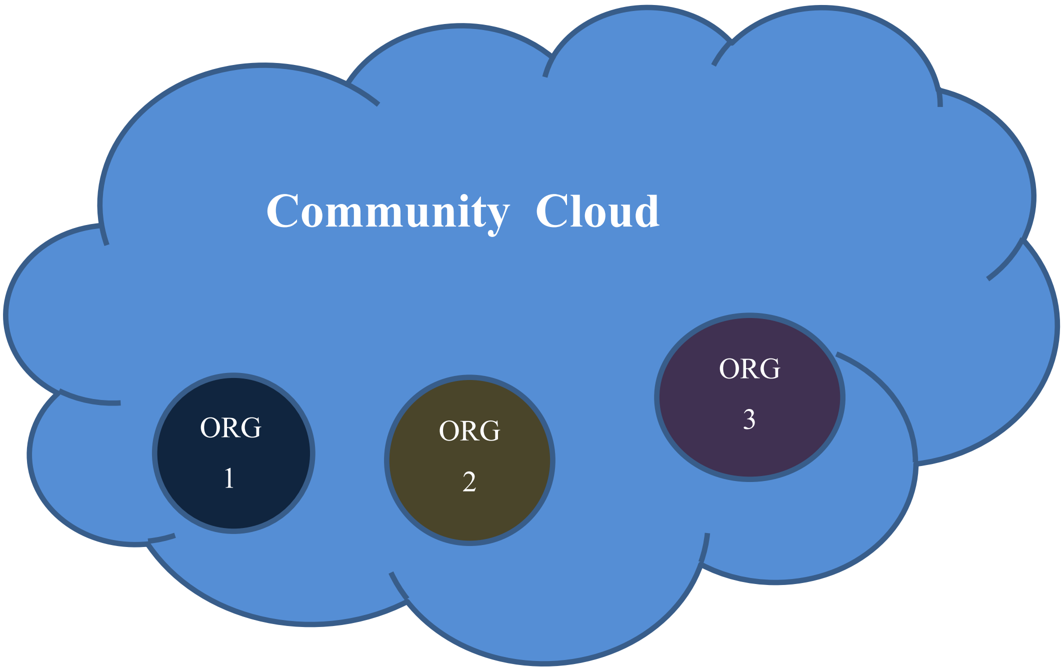

Figure 11: Community cloud Model [17]

In this model, The multi-tenant approach allows many institutions to collaborate in the same environment [18]. A Community Cloud is a decentralized infrastructure that combines services from multiple types of cloud technologies to meet specific industry needs. This model operates well for institutions and organizations with similar functions, concerns, and regulations [19].

2.4.5 Multi-cloud

An approach to cloud computing is known as multi-cloud and it is a combination of two or more separate cloud environments. This multi-cloud can use two or more cloud services, allowing businesses to avoid vendor lock-in and reliance on a single supplier [19]. Additionally, a multi-cloud approach decreases an enterprise's reliance on just one cloud hosting provider, allowing for greater flexibility in the usage of cloud services [20]. Figure 12 depicts the structure of the multi-cloud model.

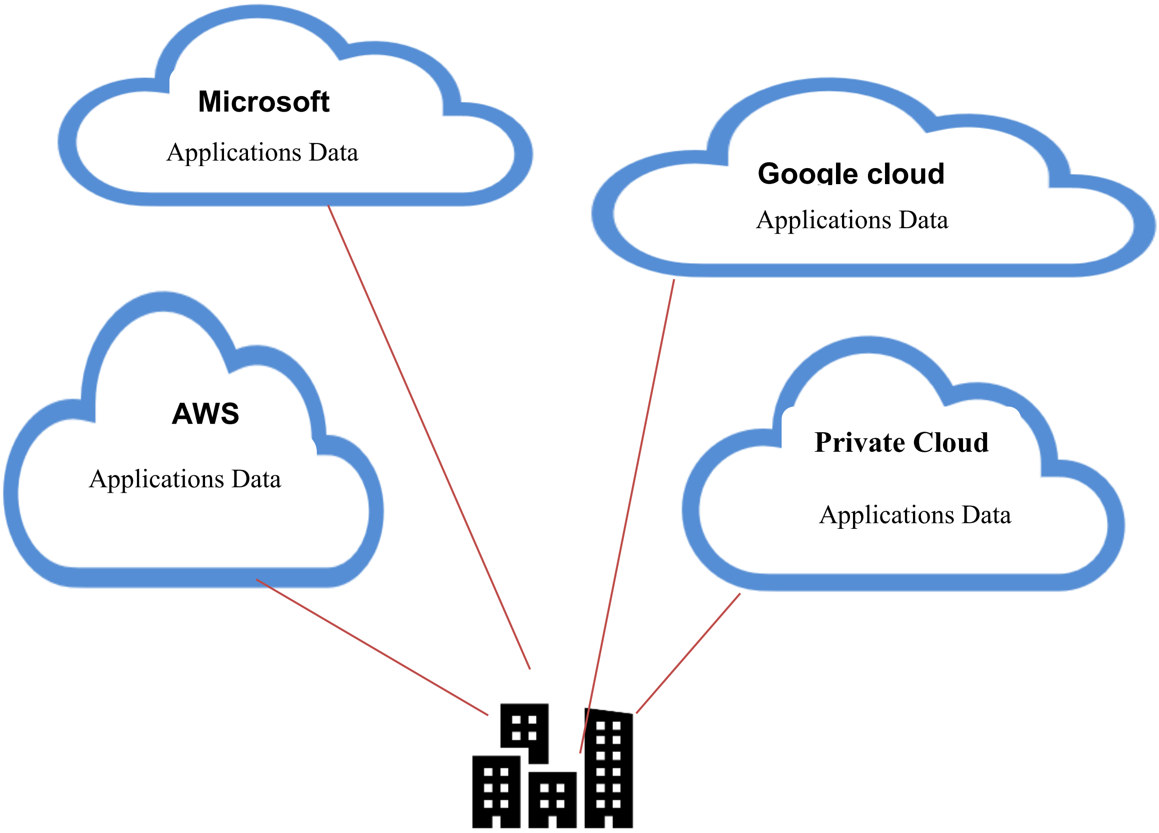

Figure 12: Multi-cloud Model [17]

Multi-cloud services are used by enterprises to redistribute computing assets and lower the risk of outages and data loss. A multi-cloud strategy allows enterprises to choose various cloud providers [17]. The key obstacle to multi-cloud security is uniformly securing data across a diversity of cloud operators. When a corporation employs a multi-cloud strategy, third-party partners manage various parts of security. As a result, it is critical in cloud implementation to correctly define and share protection obligations among the stakeholders [19] [20].

2.5 Cloud storage architecture

Cloud storage architecture represents the interaction of the elements that comprise a cloud services platform. Cloud storage architecture is largely focused on supplying on-demand storage in a multi-tenant, high-density environment [21]. On-premises, cloud, software, and middleware are the four infrastructure divisions in which the different elements and sub-elements can be constructed. The integration of these numerous components provides the core components of any cloud storage service in various ways. Figure 13 represents the model cloud architecture.

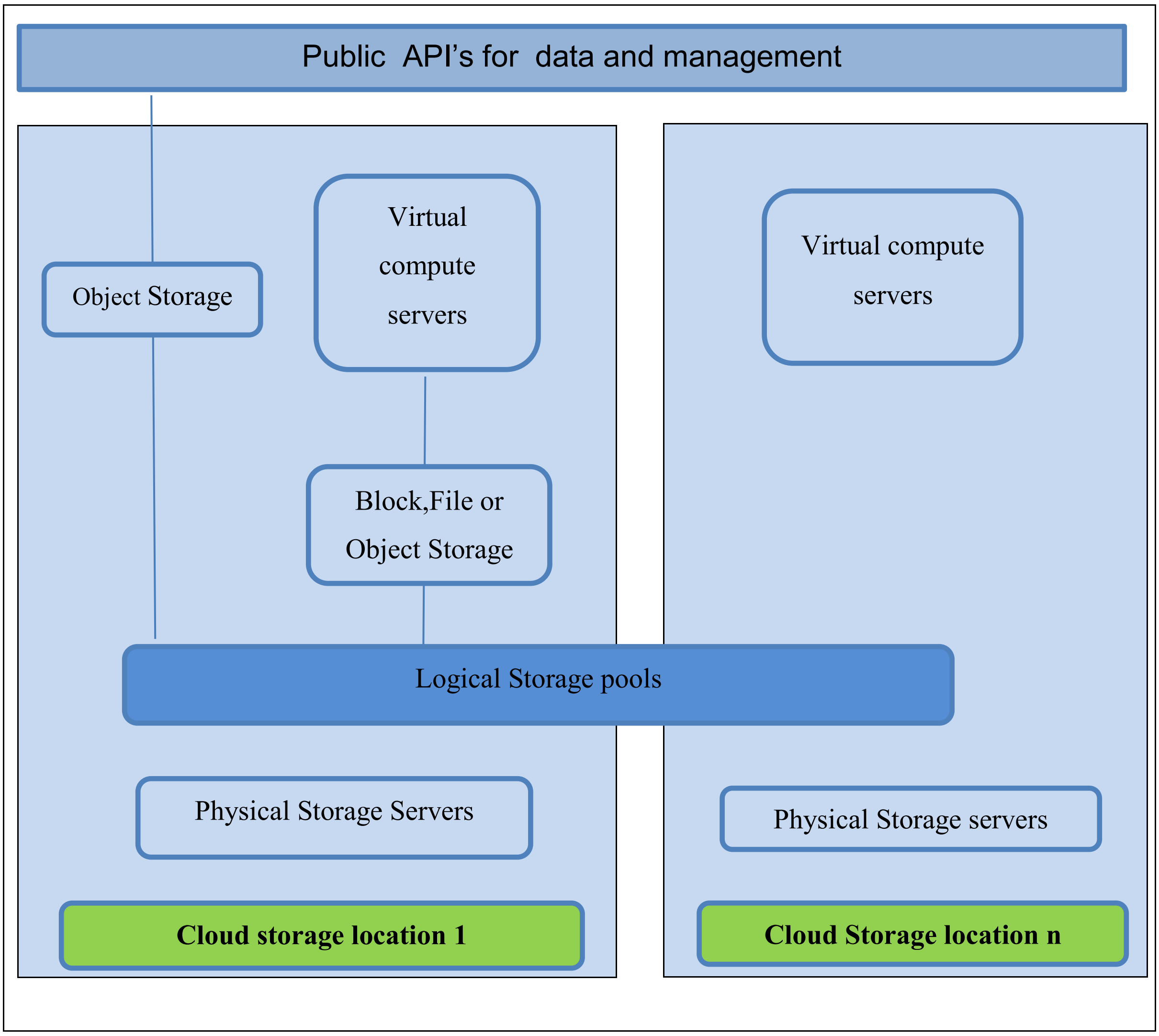

Figure 13: High levels of Cloud storage Architecture[22]

In cloud storage designs, a front end is available that exports an interface for connecting with the backend storage. At this layer, users can use web applications, file-based Internet SCSI, or iSCSI frontend. This layer is the user's first point of contact with the operator. To access the services, users must first log in with their credentials [23][22]. The basic cloud storage architecture is demonstrated in figure 14.

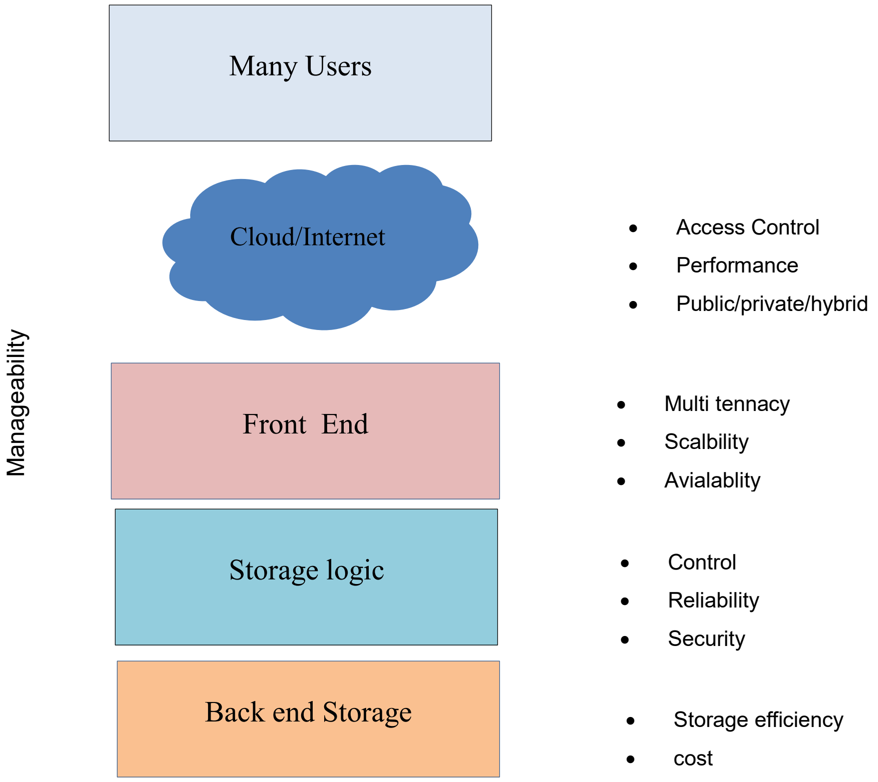

Figure 14; Generic Cloud storage Architecture [21]

The Storage logic or controller layer provides linking the front-end API to the back-end storage. Virtualization, replication, and geo-graphical data placement methods are just a few of the characteristics available in this layer. The storage logic layer also provides reliability and security to the system. The information is saved in the final layer (the backend storage). This could be a centralized protocol that manages individual applications or a typical disk backend [22]. Figure 15 depicts the four levels that make up the architecture of cloud storage.

.

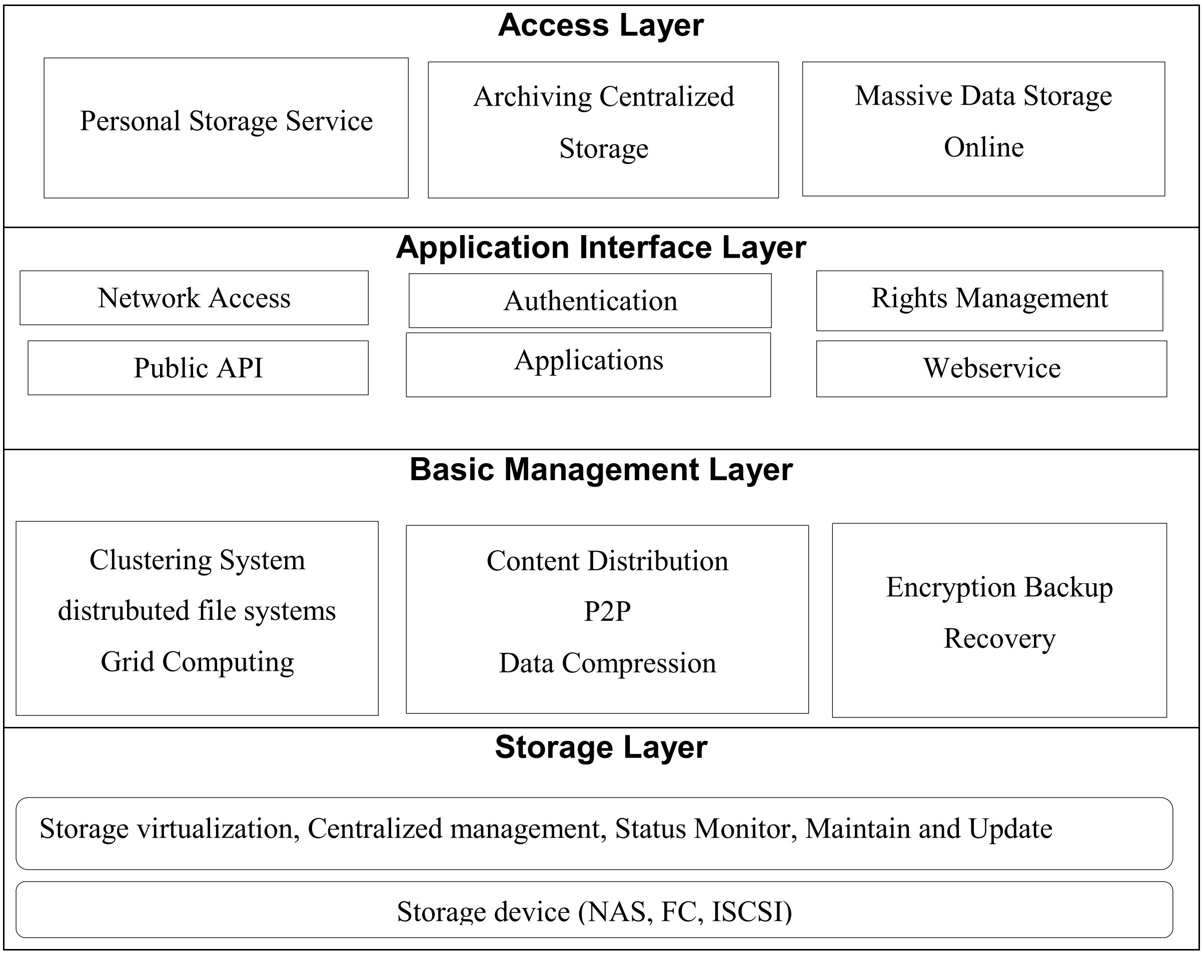
Access Layer
Personal Storage Service
Archiving Centralized Storage
Massive Data Storage Online
Application Interface Layer
Network Access
Authentication
Rights Management
Public API
Applications
Webservice
Basic Management Layer
Clustering System distrubuted file systems Grid Computing
Content Distribution P2P Data Compression
Encryption Backup Recovery
Storage Layer
Storage virtualization, Centralized management, Status Monitor, Maintain and Update
Storage device (NAS, FC, ISCSI)

Figure 15: Layers of Cloud Architecture [22]

The layers of cloud architecture have been described as follows.

Storage layer: The most fundamental part of cloud storage overall design is the storage layer, which is at the very bottom of the complete solution architecture. Fibre-channel, NAS, and IP storage devices are all examples of FC storage devices, allowing cloud storage to connect and manage a huge number of storage devices situated all over the world [21].

Basic management layer: The most significant aspect of cloud storage is the core management layer, which delivers uniform interfaces for various services across public administrations. User administration, security control, copy administration, strategy administration, and other typical data governance tasks are all integrated. The fundamental management layer may easily connect the underlying storage and the higher application to achieve syncing between numerous storage devices. As a result, the most complex part of the cloud storage architecture will be the core management layer.

Application interface layer: The application interface layer is the most adaptable portion of cloud storage, as well as the part that interacts directly with users. Operators of cloud storage carriers provide numerous user interfaces to meet their needs, as well as a wide range of application platforms and services including video on demand, network drives, and file storage.

Access layer: Access Control Layers As long as the user is approved, he can access the cloud storage system at any time and from any location using a standard public application interface. Depending on the provider, cloud storage offers a wide range of access tools and services.[21], [23], [22]

2.6 Security of cloud storage

The security of the cloud is a subset of the security of computers. It refers to a collection of regulations, technology, and control mechanisms that support data protection and services. Vulnerabilities influence the cloud service, directly or indirectly [24]. The integrity, availability, and confidentiality of cloud resources, as well as services provided by different layers, have been compromised, potentially raising new security concerns [23]. The goal of this section is to investigate several security concepts that will aid in a better understanding of cloud security issues. Figure 16 demonstrates the security model of cloud storage.

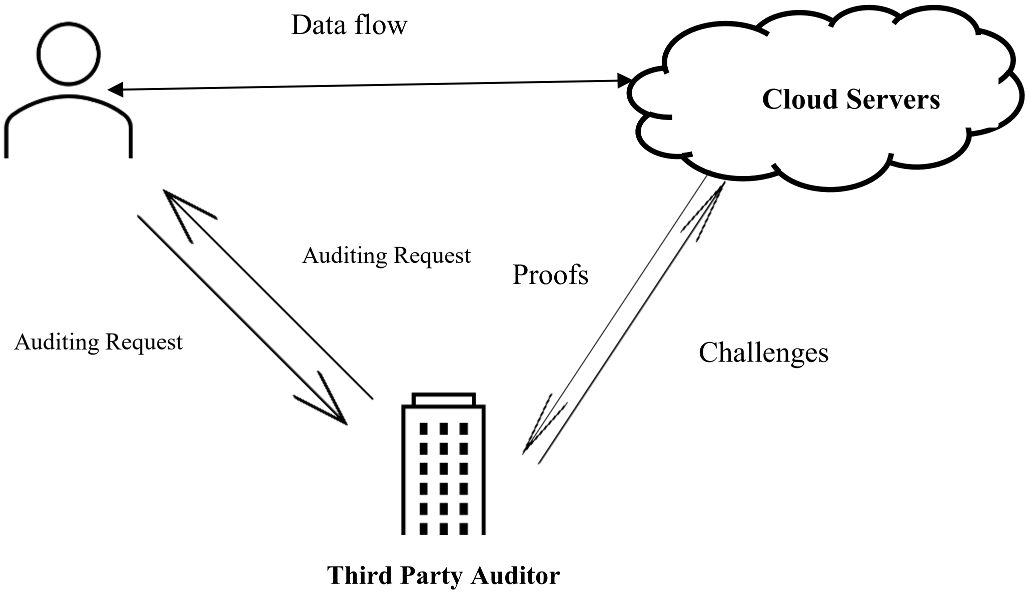

Figure 16: Cloud Storage Security Model [24]

Businesses and organizations adopt cloud services because they are inexpensive and more flexible than locally installed hardware. Storing information in the cloud, on the other hand, exposes confidential files and sensitive data to new threats because cloud-stored data is outside the reach of many of the measures in place to secure sensitive data on-premises [2]. As a result of this evolution, business organizations are growing increasingly apprehensive about cloud storage security, both in terms of IT infrastructure and information security policies.

Both cloud storage providers and businesses share cloud storage security. Companies worldwide embrace cloud services, including cloud-native development, data analytics, and machine learning to mention a few. While saving and maintaining data has never been simpler, cloud vulnerabilities are now a major concern to data security. Cloud data is typically secured using a range of approaches, technologies, and methodologies. Cloud-based solutions already have a lot of security protection built-in, which is a significant plus. This usually includes strong encryption at rest and in transport [24].

To create a geographic border and detect suspicious behavior, IP addresses and other geolocation data are used. Data-aware filtering allows businesses to keep tabs on specific situations and events, as well as who has accessed information and when. Figure 17 illustrates an overview of cloud security and access control mechanisms.

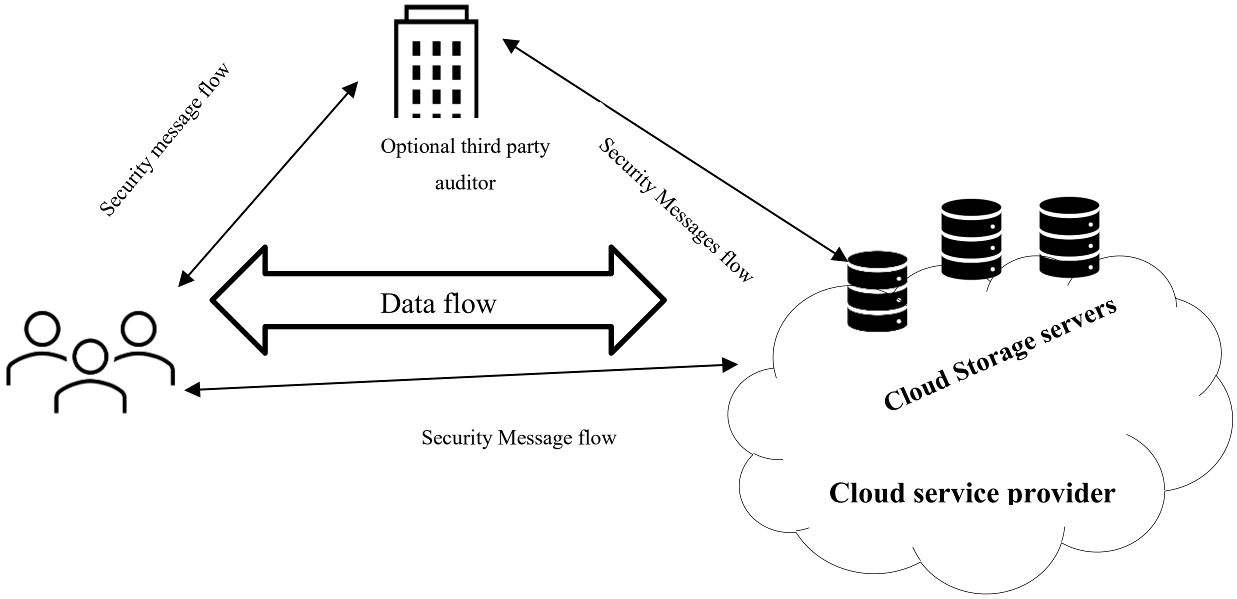

Figure 17: Cloud Security and Access Method[22]

It can be coupled with role-based authorizations and privileges. Data classification policies are used by systems to control and automate how data is kept, preserved, archived, and discarded [24]. Backup and recovery services help a company deal with not only outages but also security risks such as ransomware and data that has been fraudulently wiped. Strong disaster recovery solutions based in the cloud assure availability in any situation. Examining logs and auditing workloads can identify security flaws and risks. Cloud storage companies deploy fundamental security measures like authentication, access control, and encryption to safeguard their platforms and the data they process. Most organizations then adapt their security protocols to strengthen cloud data security and limit access to sensitive data in the cloud [24] [23].

2.7 Security challenges of Cloud storage

Security issues are increasing as cloud storage becomes more widespread. The biggest concern or challenge after a person has implemented his or her cloud-based system is security. When users choose a cloud solution from commercial cloud services, they relinquish control over physical security. Users share computational resources in a public cloud storage system. Confidentiality, integrity, and availability (CIA) [25]. Security issues, vulnerabilities, and challenges are explained in figure 18.

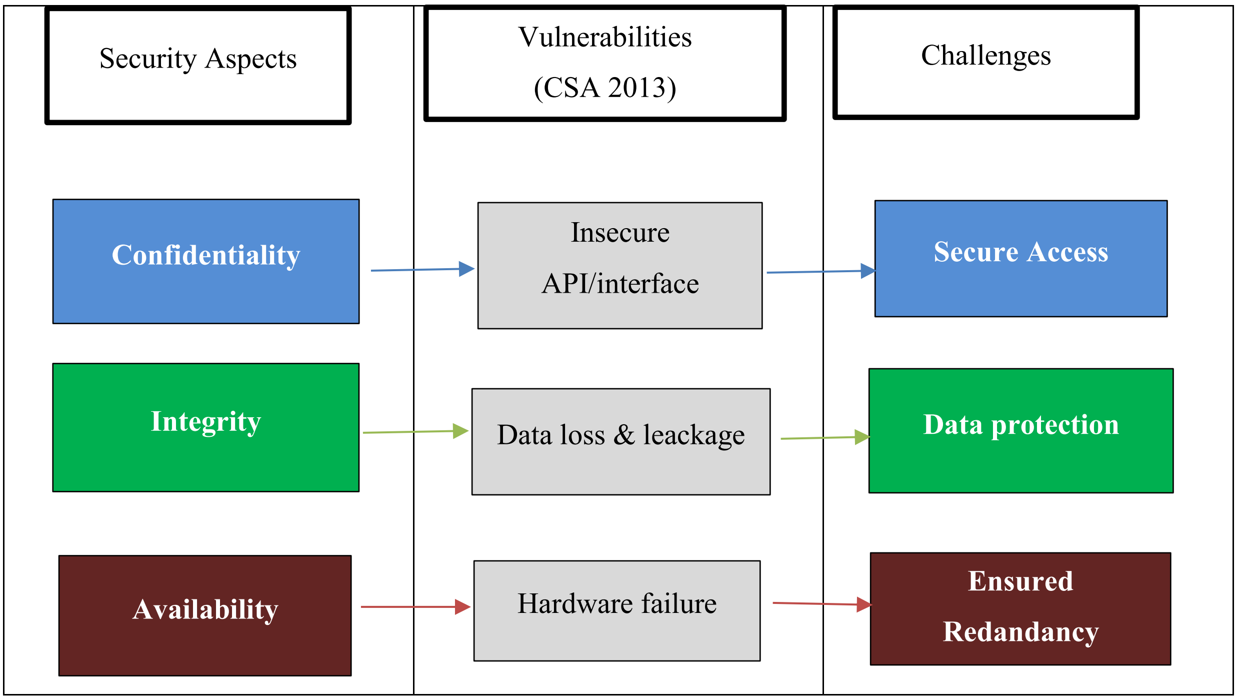

Figure 18: Cloud Storage Security Aspects, Vulnerabilities & Challenges[26]

Unauthorized access or exchange of confidential data is a significant security problem for an individual or any enterprise. Files can be stored anywhere on the globe. Information contained in any region should be compatible with cloud storage accuracy. Data privacy is eroded by exposing sensitive information, which results in a loss of data handling and exposes the company to a slew of cybersecurity risks and the legal and regulatory consequences that come with it [26]. The multi-tenancy issue makes it difficult to safeguard user data from unwanted access by other users running processes on the same physical servers. With the broad adoption of cloud storage and the fact that users are storing more essential data on the cloud, this issue must be considered appropriately [27].

Malicious third-party activities and hosting infrastructure flaws are putting data integrity at risk. Data availability is a very critical issue. Hardware malfunctions can occur at any time. Natural disasters, flooding, and even fire can all bring down systems [26]. A malicious insider is perhaps the most severe threat and the one with the most risk. A hostile insider, such as a system administrator, can gain access to potentially secret information and gain escalating levels of access to increasingly vital systems, leading to a data breach in the long run [25]. Threat actors can simply launch DDoS assaults and obtain access to critical data while remaining undiscovered by leveraging weak APIs.When an attacker employs the DoS or DDoS approach, the botnet's slaves are instructed to send fictitious traffic to the cloud, rendering legitimate cloud users' access to data, apps, or other services inaccessible [27].

One of the most dangerous aspects of multi-tenancy environments is the failure to establish impermeable isolation between tenants. Malicious actors could use the resources of some other user to get access to a company's assets or data. Multi-tenancy may increase the attack surface and lead to information leakage if the isolation measures fail. This vulnerability poses a risk to an institution's data security and privacy if it is not handled immediately [25] [27]. Attacks on the system that result in the security vulnerabilities outlined above can be carried out through three different channels. Networking, hypervisor, and computing hardware are all covered. Figure 19 represents the attack vectors and the potential security weakness in the systems.

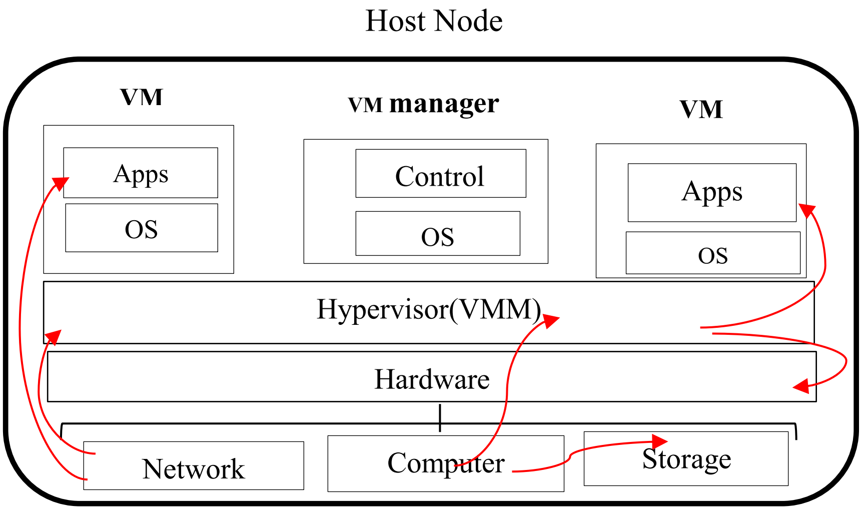

Figure 19: Cloud platform attack vectors [27]

These vectors are targeted by three sorts of attackers: internal clients, external clients, and the cloud service provider. Despite the inherent flaws of cloud storage, users have never been prevented from taking advantage of its cost savings and flexibility [25], [26], [27].

2.8 Importance of Open source cloud storage solutions

Numerous commercial cloud storage providers, such as Amazon Drive, Microsoft OneDrive, and Google Drive, provide free storage user access limitations. As a result, there is a high risk of violating security and privacy. When it comes to data backup and security, open-source cloud storage enhances efficiency and productivity. Many companies have built their cloud storage and security systems [4]. Self-hosted cloud storage is a growing trend among businesses. Sharing of information and collaboration has become more convenient due to cloud storage. Self-hosted open-source cloud storage systems are increasingly widely used for file sharing. The cloud storage area is currently under active development due to potential problems such as data loss, information hacking, and other assaults [28]. Using on-premises storage and integrated on-premises file sync and share technologies is the best way to keep data safe. Reliability, management, protection, and flexibility are all factors that support reducing the high cost of most free open-source cloud storage systems [3].

Open-source components are completely portable and can be utilized on any platform that facilitates them. The component, as well as any data linked with it, is completely under the user's control. Users do not have to be concerned about vendor lock-in, as they do with proprietary components, because of the portability and control. Users will also be able to avoid the increased danger of granting access to their data and systems to third parties. Open-source software is completely observable and quantifiable [3] [29].

One of the grounds for the popularity of open-source cloud storage is that open source cloud solutions are a community, they are more likely to be able to fix any issues and disseminate solutions faster than proprietary software competitors [4]. Most of the improvements and upgrades are free and available to everyone because of the cooperative nature of open-source software. Vendor lock is a common problem for commercial cloud storage systems. clients can switch between cloud providers without any problems while using open source cloud software compared to commercial cloud applications. Open source is also distinguished by its ability to work in various paltforms [29]. Users can alter the source code of open source cloud systems, leading to increased functionality and flexibility. For example, suppose a new update is required for improved functionality of an open-source application. In that case, it can be implemented swiftly with the assistance of an experienced developer team [4].

Users get access to a whole commonality of the field specialists who regularly create and update the software when they use open-source software. Better protection is achieved as a result of this mutual effort. Even though commercial items can be safe, they lack the perspective variety found in a broader community. Data breaches can occur in any context; the only difference is that resolving them is different in each [3].

3 Comparison: Owncloud vs. Nextcloud vs. Seafile

In this chapter, a comparative study is analyzed the three open-source cloud solutions (ownCloud, Nextcloud, and Seafile) based on their features, architecture, security mechanisms, and vulnerabilities.

3.1 Features

This section has described the significant features of the three open-source cloud storage solutions (ownCloud, Nextcloud, and Seafile). The following significant features of owncloud, nextcloud, and seafile are identified in this study, and table 1 represents a comparison view of these cloud solutions based on the features.

Table 1:Generic features of Owncloud, NextCloud and Seafile [30], [31], [32], [33]

General Features			
	ownCloud	Nextcloud	Seafile
Licensure	Open source (AGPLv3)	Open-source (APGL)	Open source
Storage capacity	Unlimited	Unlimited	10 GB
On site self-hosting	Yes	Yes	Yes
Classification	File sharing and sync, cloud storage, content collaboration	Document collaboration and management, file sharing, Team work, screen sharing, etc	File sharing
Support for large file size	Yes	Yes	No
Usability	Easy to use in general	The UI is simple to use.	required some technical skill
Scalability at global level	No	Yes	Yes
The features for enterprise	Network drive support, SharePoint integration, File firewall, Single sign-on, etc	Collabora Online Office, Custom branding Controlling the data flow, Integrated account management, etc	Remote wipe,Role-based account management,Lock files
Client apps for mobile devices	Android, IOS	Windows, Android, IOS	Android, IOS
Desktop clients	Linux,Windows, mac	Windows, mac linux	Mac windows linux
Realtime alert	No	Yes	No
Dashboard and share note	No	Yes	No

3.1.1 Features of OwnCloud

Open-source file hosting service ownCloud was founded in 2010 to simplify the method of building and deploying cloud-based storage services. OwnCloud is an exquisite alternative for anyone curious about hosting their cloud storage system. OwnCloud was the first enormous alternative to commercial platforms like Dropbox, and it did so in an equally user-friendly package. Over the years, OwnCloud has grown and expanded into a major project supported by a large organization. It improved its user interface and made apps available for nearly every mobile and desktop platform [30]. It also works well for enterprises and entrepreneurs who choose to establish a cloud service for widespread public use because it does not require a high level of technical expertise. As shown in Figure 20, ownCloud offers consumer-grade usability and maintains file sync and share.

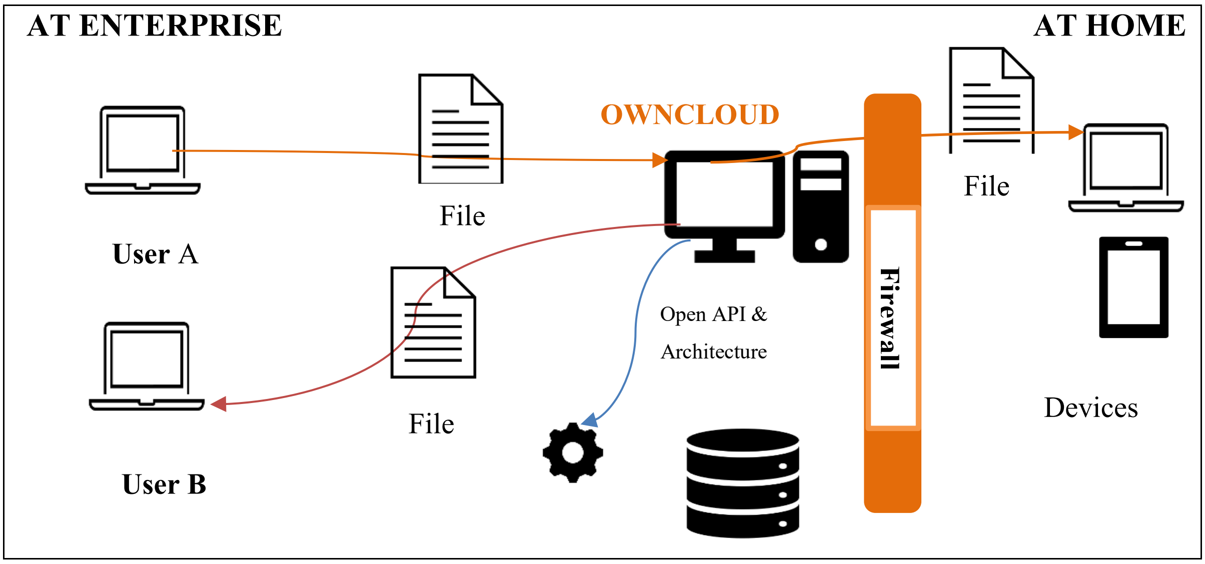

Figure 20: The functionalities of ownCloud [34]

Clients are often authorized to use OAuth, an open industry-standard protocol. It considerably boosts security while making it easier to integrate third-party apps or online services [30][35]. Organizations can use the audit functionality to track what users and admins do with which data, preventing misuse and conducting compliance audits. It's especially helpful in areas where data access is restricted, like within the health and financial industries [30] [36]. The integration of ownCloud with eM Client enables a protected and straightforward method of transferring documents over email without the utilization of a third-party service. This approach, additionally to being simpler to operate, provides better levels of protection for both the sender and also the recipient. If the document is later destroyed, the link will not function properly. Furthermore, it's possible to guard the files with passcodes and set expiration dates for them [30] [35].

It is executable to automate the rules-based archiving and deletion of files using the ownCloud file lifecycle management app to suit a range of standards and requirements. The files firewall compares each access request against a group of rules and restrictions. It enables information technology departments to limit file access to certain user groups outside of company headquarters, for instance. Additionally, access to files with specified tags and from high-risk countries is often restricted [30]. The facial recognization verification method is conducted via the biometrically-secured authentication techniques in the particular operating platforms enabled by the ownCloud Apps for the platform(iOS and android). A seamless link between ownCloud and Microsoft Teams is enabled by combining the two apps, allowing for easier control and improved security.

The metrics app allows for the focused gathering of important evaluation metrics, resulting in data for reporting [30]. The comment feature simplifies and speeds up coordination. The Outlook plugin allows the client to deliver emails with attachments that are automatically saved in their ownCloud and transmitted as a URL rather than enclosed with an email. It is increased security and allows users to upload documents of any size. Public URLs allow users to share documents and folders with friends and family who do not have an account. Passwords and expiration dates can help keep track of personal data.[30][36]

Data is protected using cutting-edge cryptographic techniques. ownCloud's modular and flexible encryption architecture allows for unique configurations to meet any threat level or regulatory requirement. Users may effortlessly link their ownCloud to Microsoft Teams, allowing them to share information and files consistently and safely across Microsoft's workstream collaboration tools. The Activity Stream highlights what was done with which document to others who used the same document and directories [30][36]. The True Secure View functionality allows clients to transfer confidential documents securely while maintaining high levels of control. Document classification automatically secures files based on content and tags. Enabling the virtual files feature allows users to sync a virtual file rather than the entire file, downloading the actual thing only when needed. It reduces the amount of data stored and the amount of bandwidth used [30]. Owncloud's architecture allows users to save and share content of any size.OwnCloud provides users with customized access to their data. Using a full-text search, they can get to the file they want faster [36].

3.1.2 Features of Nextcloud

Another popular open-source cloud storage solution is Nextcloud, developed in 2016. NextCloud is a fork of OwnCloud. Additional features, such as full cryptographic algorithms, are currently available in Nextcloud. The group then started working with well-known open-source software projects like LibreOffice. This collaboration resulted in Collabora Online Office, an open-source replacement for Google Docs. Nextcloud is constantly expanding, solidifying its position as the most dependable and robust open-source cloud storage solution [31] [37].

Teams can securely connect via video chat while also sharing documents and sending email messages, one of Nextcloud's key features. The functionalities include remote access, user authentication, compliance management, and audit tracking [31][38]. Safety pen-tests and external consultants have demonstrated that nextcloud was built with proper security measures[31]. Login security is supported by artificial intelligence (AI), two-factor verification, brute force prevention, and industry-leading features like video verification, end-to-end, and server-side encryption [37].

Nextcloud is one of the prominent open source file synchronization and collaboration system [31]. It offers the foremost comprehensive range of incorporated functionalities and interfaces available. There are quite 200 "apps" to choose from, each with enhanced security, collaboration tools, and infrastructure interfaces [39]. Users may utilize Nextcloud Flow to automate repetitive procedures and improve business processes. Edit office documents collaboratively with colleagues and takes notes while on a video chat [31] [39].

As a part of its compliance efforts, Nextcloud adheres to industry standards like Clause 14 of ISO/IEC27001-2013 yet as associated standards, advice, and security concepts [38]. Powerful features often cause complicated user interfaces, which might influence productivity and make it hard for organizations to quickly adopt new technology. Nextcloud does not limit users' abilities to avoid overwhelming them [37]. File locking helps users avoid disputes with colleagues who try to change the identical files at the identical time as they're locking the files [31]. While some items, like office documents and notes, could also be changed in real-time within the browser, others must be downloaded to be modified. Collisions are prevented if the file is secured. If other users have questions about what's occurring, they will quickly contact other users in chat or through a remark [38].

Nextcloud incorporates a simple-to-use collaborative note-taking tool called Nextcloud Text. It provides a variety of rich text formatting options, including headers, bold, italics, pictures, and links. During a bunch of editing sessions, author colors are accustomed to indicate who made which changes[31] [37]. Under the GDPR and the CCPA regulations, businesses are held liable for any breaches of user privacy or failure to adhere to a high degree of data protection. By simplifying regulations, the open-source cloud storage solution Nextcloud lowers business costs and risks [31]. At healthcare service and biomedical facilities, patient safety is a top responsibility. Nextcloud provides patient information accessible to healthcare providers when they need it, with the highest level of assurance, protection, and confidentiality at a fair cost, through a simple implementation [31].

Utilizing WebRTC search engine techniques, Nextcloud includes an option for embedded audio/video chat. The deployment of a STUN/TURN dedicated server significantly improves its option to communicate with clients in severely firewalled environments [31]. The template support in Nextcloud allows users to quickly create a spread of files. within the templates folder, users can keep track of their templates. There are a variety of pre-made themes and plugins in the nextcloud system [31][37]. Nextcloud includes extensive keyboard ease of access and screen reader assistance to ensure that those with impaired vision can use it [31]. Collabora Online is a robust online office suite based on LibreOffice that supports collaborative editing in all popular browsers [38]. Account management is embedded into Nextcloud, along with an additional 2FA (two-factor authentication), enabling it simple to generate and alter accounts. The Monitoring app allows administrators to keep track of a Nextcloud system's activity and performance [37].

3.1.3 Features of Seafile

Seafile is an open-source project that is developed in 2012 and has a long history of dependability and security. It has developed a tremendous rise, and it has been sponsored by a company that provides Seafile with enterprise assistance [32]. It has progressed from a peer-to-peer file-sharing tool to a full-featured competitor to Dropbox and Google Drive over time. Among the features are file encryption, public link collaboration, automated cross-platform synchronizing, and per-folder controls [32] [40].

Seafile is a very robust and efficient open source file sync and sharing platform. Using this service, anybody may save files on a single server across several platforms [40]. Using Markdown format, users can put Wiki material straight into a library. Besides Wiki papers, other types of data can be housed in a library alongside them. Because of this seamless connection between Wiki and Cloud Storage, the conventional Wiki system's attachment size limit is no longer a concern for users. An HTML5-based Markdown editor, such as the WYSIWYG Markdown editor, makes it simple to edit Wiki pages in any web browser. Comprehensive knowledge management tools like full-text search, file tagging, relevant documents, and document review are available to help users be more productive [32] [40].

File locking is provided by Seafile to prevent multiple users from making modifications to the same document at the same time, which can escalate to avoid conflicts. Documents can be locked in the web UI or on desktop clients, depending on the user's desire. When an office file is opened, it is automatically locked [32]. Seafile preserves versions of documents and snapshots of folders in its database. Users may rapidly restore a file or folder to a prior version by simply clicking on the Restore button. Snapshots for folders are a practical approach to secure files from ransomware threats [32]. Using deduplication technology, document versions may be maintained effectively while utilizing less storage space. Incorporating Seafile's WebDAV interface with various mobile applications, such as DocumentsGoodReader, and permitting them to see files is feasible [40].

A REST API is also provided by Seafile for collaborating with third-party programs using HTTPS Seafile manages backups using mysqldump and rsync. Additionally, it enables real-time backups via the implementation of a dedicated server to obtain backups from the primary server instance [32]. Seafile integrates with MS Office Online Server to provide interactive collaboration and co-creator for Microsoft office files. Seafile also includes a preview function for films, audios, PDF documents, pictures, and text documents [32]. Seafile contains the self-fsck function, which assists administrators in recognizing as well as deleting misleading content, fixing faulty libraries, and exporting all information in the event of a system crash [40].

Seafile offers free software downloads, allowing individuals or businesses to create their private file synchronization services. Users can utilize numerous synchronization servers and switch between them with Seafile. Seafile has group abilities, as a result, users can form and connect groups, as well as share files among them securely. Seafile uses a data format and distributed synchronization technologies that are related to GIT to improve dynamic synchronization and huge file management [32] [40]. Seafile provides a driving client that allows users to save data locally on their workstations. The drive client uses storage space on the Seafile server to give customers greater local storage. The documents in the drive client are not synced with the rest of the system[32].

3.2 Architecture

This section describes the architecture of each of the three cloud storage systems that were chosen for comparative studies.

3.2.1 Architecture of ownCloud

The key component of the ownCloud system is the ownCloud server. The server offers a secure internet interface via which admins may manage their ownCloud's entire systems. Legitimate members may allow and disallow features, define rules, construct archives, and monitor individuals using this interface [34]. Figure 21 shows the overview of the ownCloud solution architecture.

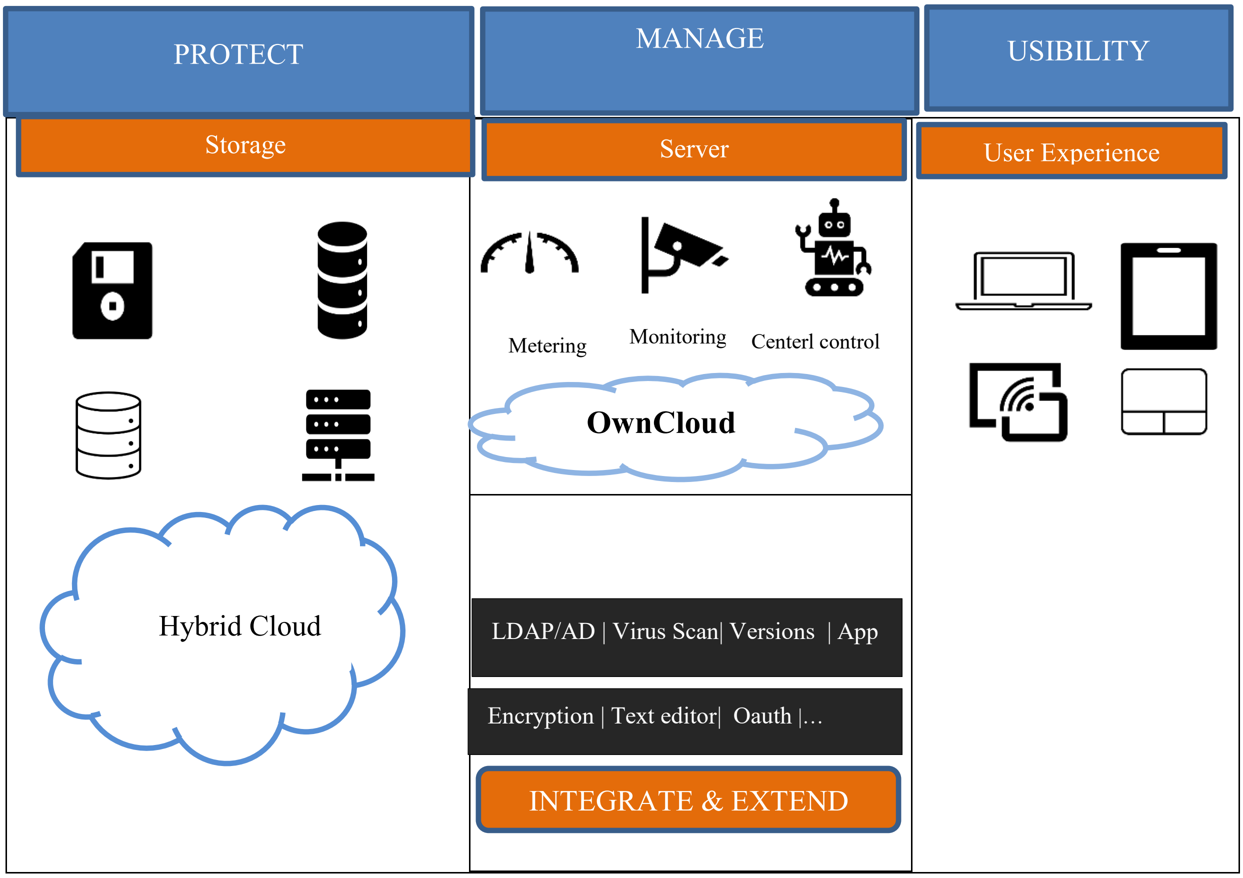

Figure 21: Solution Architecture of ownCloud [34]

The server supports and protects API access to ownCloud, while also serving as the primary processing engine that allows the service to offer reliable and excellent file-sharing options. Figure 22 illustrates the core server architecture of ownCloud.OwnCloud maintains user information in industry-leading file formats and is therefore compatible with the majority of industrial file formats.

Individual features like a web word processor, document versions, virus tracking, and server-side protection are all incorporated in the main software of ownCloud [34]. There are numerous extra features that users can develop, from media streaming to calendar and contact syncing to customize verification systems and API-based storage [41].ownCloud interfaces effectively connect with all current components and services, from traditional data backup and detection of breaches to log management and Data Loss Prevention (DLP) solutions. On-site storage may simply be mounted on the ownCloud server in most deployment settings, such as the mount point /data/storage device [34].

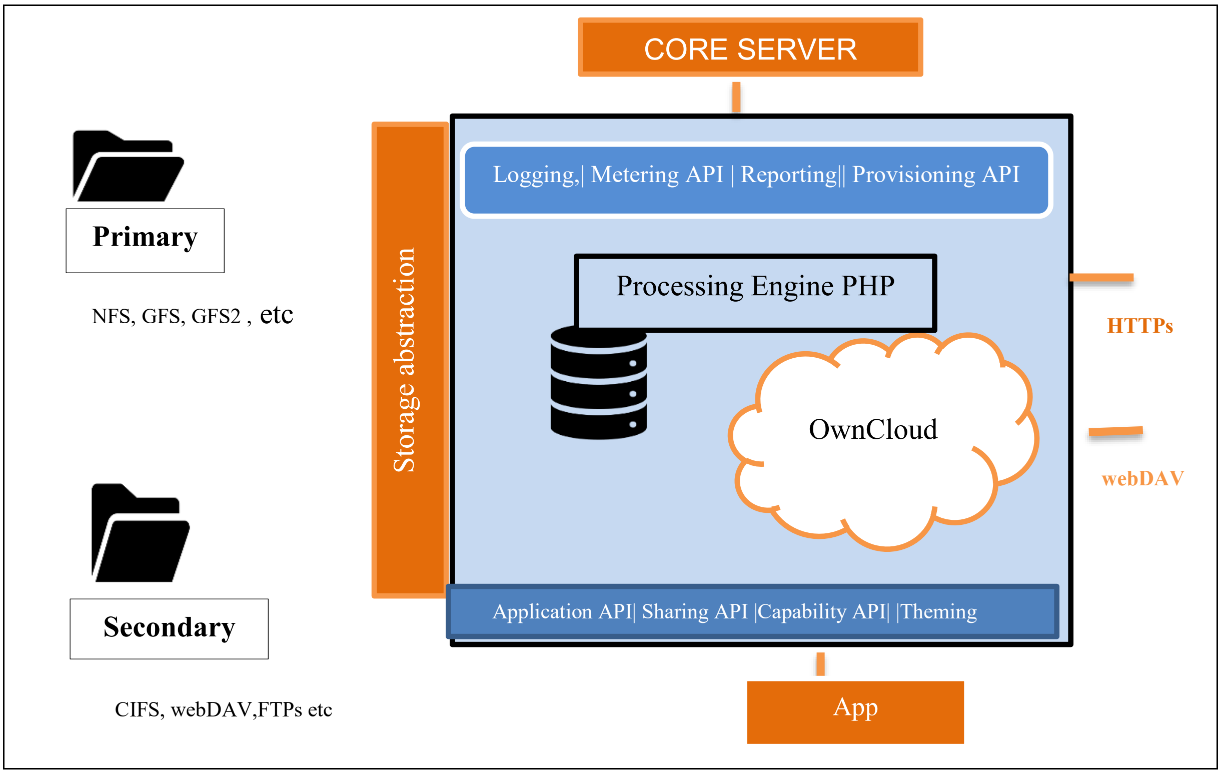

Figure 22: ownCloud Server Architecture [34]

Business directory integration and file "firewalls" give administrators a tremendous amount of freedom and control. ownCloud may be configured to use the included encryption module, which provides an extra degree of protection to user files while they are in transit [41]. The deployment model of ownCloud is described in figure 23.

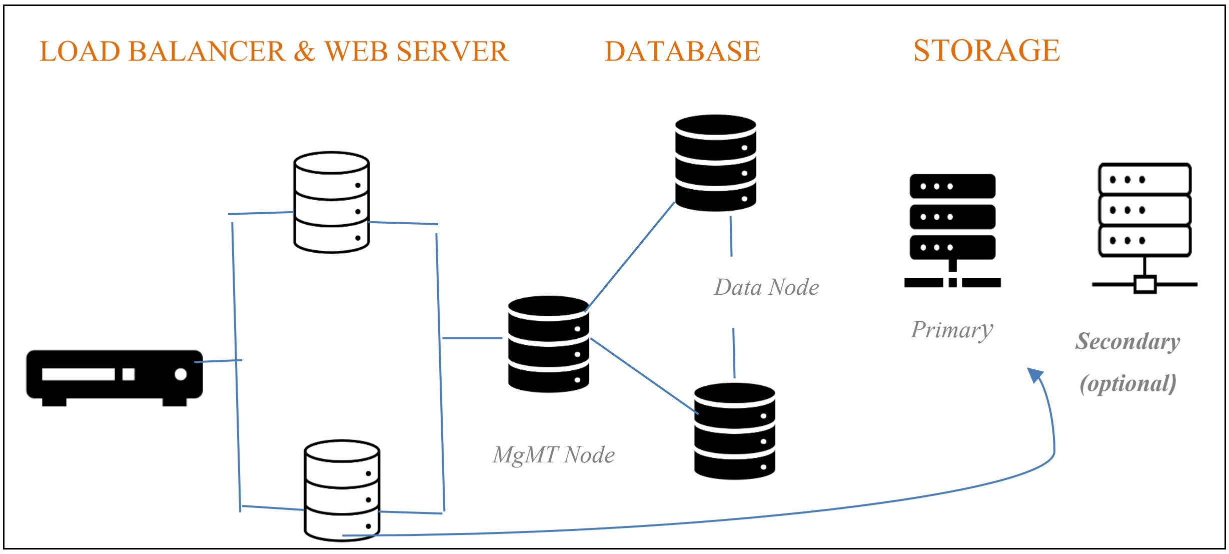

Figure 23: Common Deployment Architecture for OwnCloud [34]

Integration with existing technology is made available using ownCloud's plug-in applications, and Integration plugins deliver characteristics for instance AD (Active Directory) and LDAP (Lightweight Directory Access Protocol) interface for member profile creation and verification, among other things. Mobile libraries, open APIs, and plug-in apps are all available for use in ownCloud to help with the development of customized integrations [34].

OwnCloud is typically implemented as a multi-tier load-balanced web app in a cloud server or integrated cloud architecture. OwnCloud may be installed on conventional, virtual, or private cloud servers utilizing a digital device footprint. A load balancer is often linked to multiple web servers on the deployment's forepart. The PHP script is served on the web servers of its cloud. The virtualized file cache, shared file lists, client and group metadata, and storage that allowed ownCloud programs require are all then attached to a database for user data. All servers also are linked to a shared back-end storage system, which is commonly a distributed filesystem [34] [41].

3.2.2 Architecture of Nextcloud

The server of Nextcloud is a PHP web-based software component (application) that runs on a Linux web server such as NGINX or Apache. It uses a database to store and exchange files, user information, program files, and settings, as well as file metadata. Nextcloud provides support for the databases MySQL, MariaDB, and PostgreSQL. As a performance measure, a

REDIS caching server may be utilized to accelerate data access while also reducing the burden on the database [42]. Figure 24 shows the diagram of a server architecture of Nexcloud and an illustration of a Nextcloud system with an NFS drive layer, Servers for REDIS CACHING, an LDAP member directory, multiple db (database) servers, and a traffic shaping (load balancer).

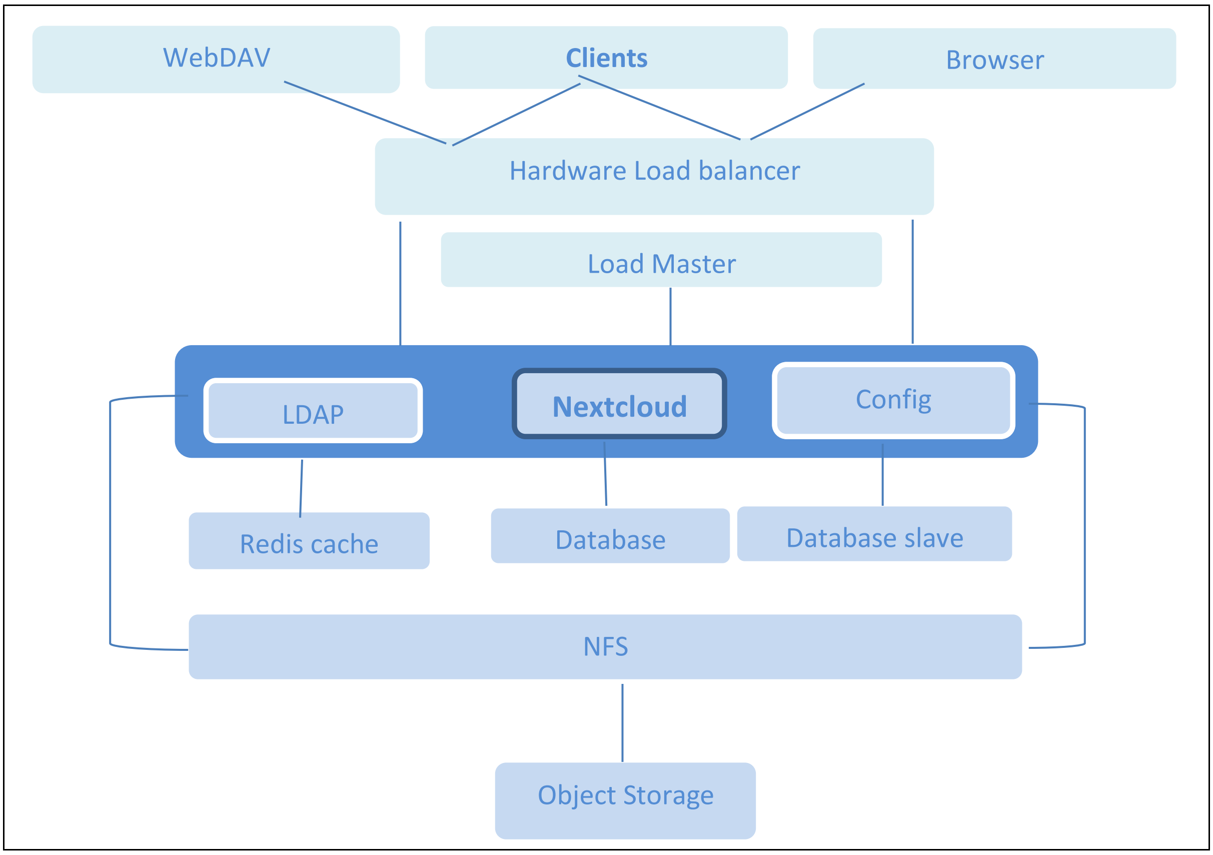

Figure 24: NextCloud Server Architecture [43]

If a server can be equipped with a storage protocol like NFS or GFS2, the storage layer can use that protocol. This includes Red Hat Drive, Windows Network Storage, IBM Elastic Storage, and object stores that can be utilized with SWIFT and S3 [42]. The system is set up to automatically allocate storage space based on client directory listings, which allows for data separation and multi-tenant implementations. The system can be set up to automatically assign storage space based on client directory entries, which allows for data separation and multi-tenant implementations. Administrators have a great lot of control over who can access and share their data with Nextcloud [42] [43]. Administrators can match storage services to users based on group participation or other criteria by leveraging user data from LDAP or Active Directory. The storage path identifier can be acquired after joining LDAP/AD, allowing users to be directed to different storage paths [42] [44].

Aside from primary storage, the system administrator can use the settings panel to mount and install a variety of alternative storage systems that can be differentiated for specific groups or clients. This allows a subgroup to access a CIFS or FTP system while some other clients maintain control over their private dropbox [43].

Consumers get a familiar and easy-to-use experience via Web browsers, iOS, and Android applications, and desktop synchronization tools, while administrators have a lot of control over data access and sharing. They can also use a range of collaboration tools, such as Nextcloud's WebDAV compatibility to access their files. Nextcloud's architecture is flexible, allowing administrators to add and remove functionalities as needed. REDIS caching can improve efficiency on deployments with tens to hundreds of thousands of clients. Following this stage, clients use the Federation to allow additional extensions to millions of clients [43]. Nextcloud comes with several server APIs that make it possible to combine it with other systems. Nextcloud's open-by-design architecture gives IT unprecedented freedom while effortlessly integrating with current legacy infrastructure. As a consequence, a cost-effective solution is designed that gives IT complete control over company data while providing end-users with the appealing, productive layout they expect [42].

3.2.3 Architecture of Seafile

The architecture of Seafile cloud solutions is described as follows. A three-tier architecture is used in the Seafile cluster solution. These tiers have been explained as follows.

Load balancer tier: This tier is responsible for routing data flow to the servers. Many load balancer nodes can be installed to achieve high availability.

Backend storage: A shared storage cluster, such as S3, Swift, or Ceph, serves as the backend storage.

Seafile server cluster: In a cluster, there are several Seafile machines. If a balancer detects a failed instance, it will stop delivering data flow to it [45]. Figure 25 depicts the architecture of the sealife cluster.

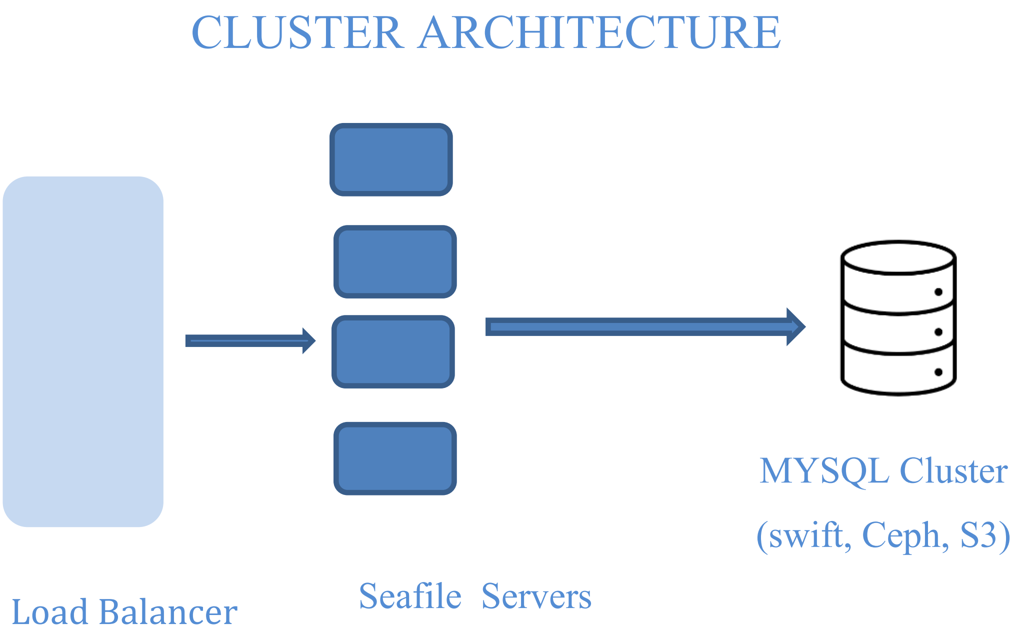

Figure 25: Architecture of Seafile cluster [45]

Seafile is a horizontally scalable, stateless file server. MYSQL Cluster contains the user-library mapping as well as the commit ID. All information and metadata are kept in object storage [46]. Figure 26 describes the components of the architecture of Seafile's server.

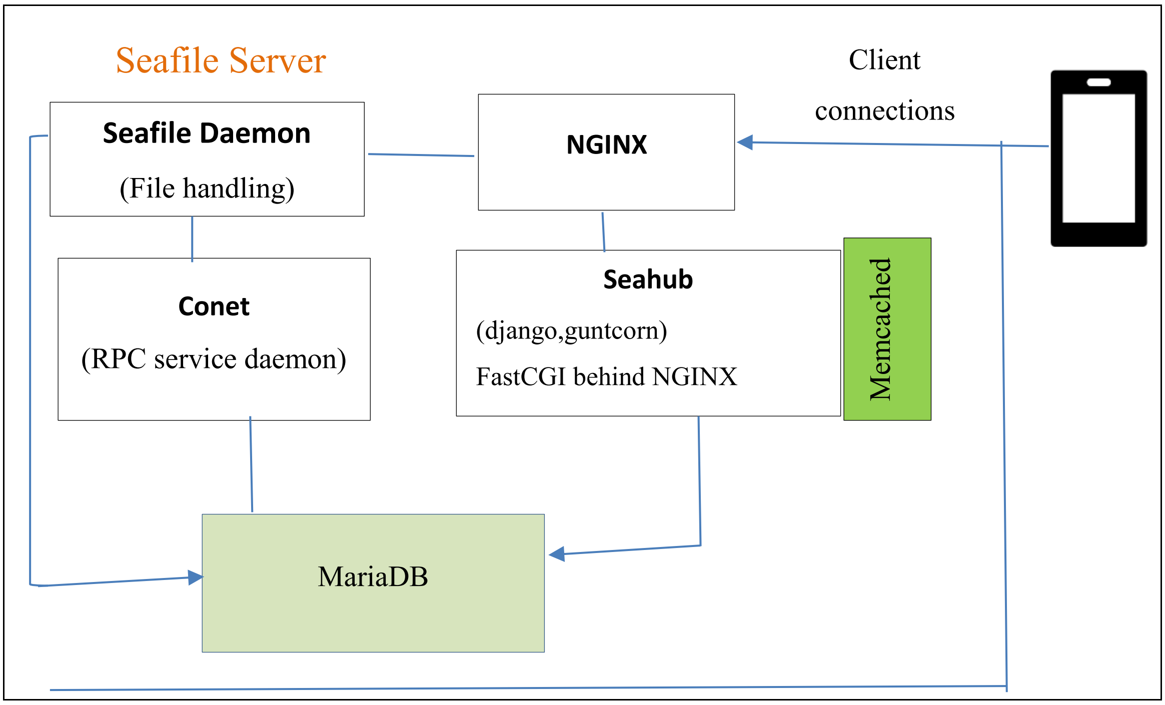

Figure 26: Seafile Server Architecture [46]

The two major elements of the Seafile server links are the webserver (Nginx/Apache) and the application server. Client requests are delivered to the Seafile application server through the webserver. In the event of system failure, the load balancer is liable for recognizing and rerouting queries. The overview of the Seafile architecture has been shown in figure 27.

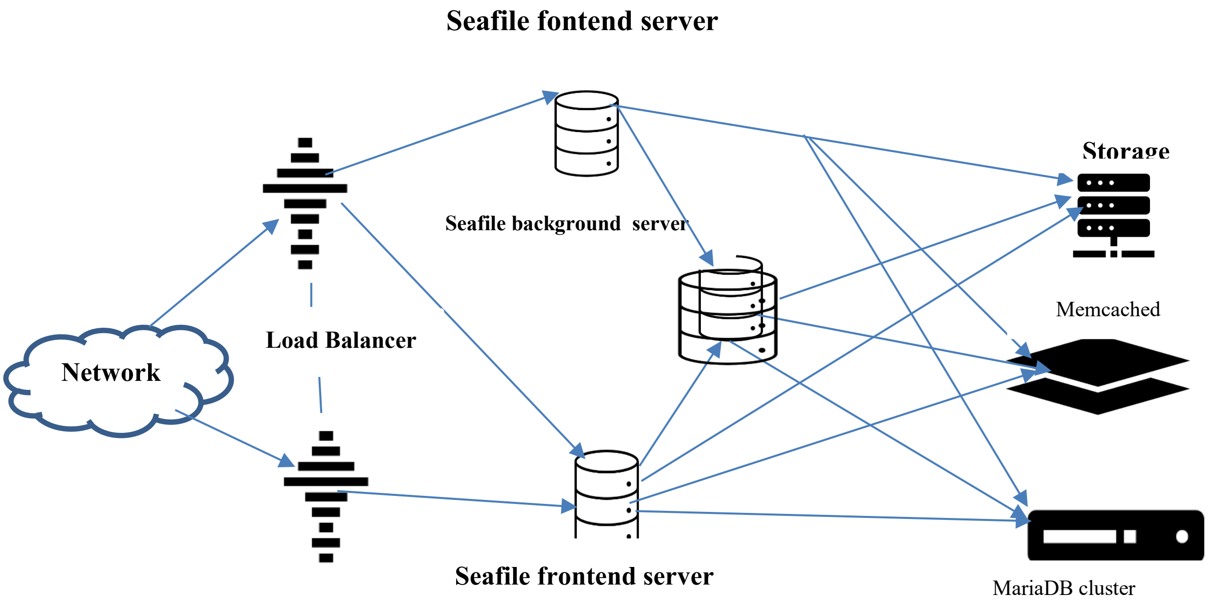

Figure 27: Overview of Seafile Architecture [46]

The backend (background) server of Seafile is in charge of full-text indexing, malware detection, office document reading, and LDAP synchronizing. It is recommended that it should be operated on a devoted server for best results. The application servers of Seafile are different and distinct in terms of operation [45] [46].

The data from the user is split into two divisions: the first is kept in a MySQL database, while the other is maintained on a backend storage cluster (S3, Ceph, etc.) [45]. Users receive an equal level of performance from all application servers. To perform effectively, all servers (applications) must be linked to a similar db (database) or db cluster [46] [47]. Figure 28 illustrates the client's access technique and process of Seafile.

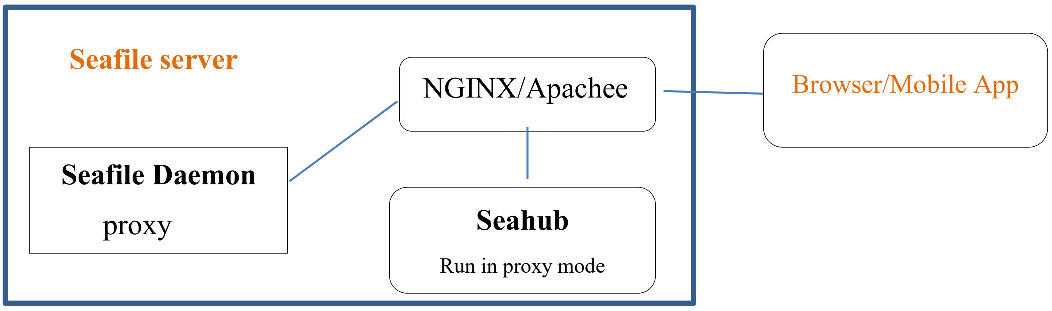

Figure 28: Process of Seafile clients' access [45]

Seahub: Gunicorn, a lightweight Python HTTP server is included in the Seafile server package, is used to serve the website. Gunicorn comes pre-installed as a unicorn application. Nginx or Apache in WSGI mode can also be used with Seahub.

Seafile server: It is responsible for raw file upload, download, and syncing. Data flow can be diverted to the local 8082 port by configuring Nginx/Apache.

The connectivity to the service of Sealife may be implemented beneath a web server. HTTPS can be used to secure all data traffic to and from the service.[45] [46]

3.3 Security and Encryption Techniques

One significant challenge when preserving data in the cloud is security. Thus, before selecting any cloud storage solution, individuals or enterprises should be aware that the cloud solution must have significant security techniques and various security measures to protect data against malicious threats [48]. This section has studied the key security aspects of the three open-source cloud storage solutions, which may help in selecting the expected cloud storage solution. In this research, the following significant security aspects of owncloud, nextcloud, and seafile were discovered, and table 2 represents a comparison view of these cloud solutions based on the security specifications.

Table 2: Security features of ownCloud,Nextcloud and Seafile [33], [49], [50], [51]

Security			
	ownCloud	Nextcloud	Seafile
Encryption on the server-side	Yes	Yes	Yes
Encryption on the client-side	No	Yes	Yes
Audit log	Yes	Yes	Only in Pro Edition
Binding 2FA	Only OTP	SMS/OTP/U2F/Telegram alert	OTP/SMS/U2F
Privacy links	Yes	Yes	Yes
Protection against brute force hacking	Yes	Yes	Yes
Conditions of service	No	Yes	Yes
Suspected login detection using machine learning	No	Yes	No
Integrated information request or account cancellation	No	Yes	No
Native SAML	No	Yes	Yes
NIST standard password security	No	Yes	NO
Auth via ENV parameter	Yes	Yes	Yes
AD/LDAP	Yes	Yes	No
Kerberos	Yes	Yes	Yes
Regulate over file access	yes	Yes	Yes
Privileges to app connectivity	No	Yes	Yes
WEB User interface Protected WITH CSP 3.0	No	Yes	No

3.3.1 Security and Encryption Techniques of ownCloud

ownCloud offer many security mechanisms to secure data storage including files. It ciphers data both in travel and at residue. ownCloud is one of the safest applications for sensitive data because of its Privacy-by-Design and Zero-Knowledge-Architecture and available on-premises hosting. It authenticates users using a variety of different authentication factors and integrates with current identity providers [52]. Permission management and policy enforcement are rigorous, and robust monitoring ensures that access is traceable and confined to authorized users [30].OwnCloud encrypts data in two aspects: when it uses the HTTPS protocol to send and receive data through the server, and when it stores data on a remote server using the encryption app. The Secure Socket Layer (SSL) protocol is applied to safeguard communications between web servers and browsers. Certificate Authorities, an impartial third party, can be used to determine the identities of either one or both parties involved in a transaction (CA) [48] [52]. The following are some important security features.

2FA: Owncloud uses the 2FA method. It is more difficult for an intruder to get entry to a system using a legitimate user's credentials if two-factor authentication (2FA) and multi-factor authentication (MFA) are used in conjunction [52].

Firewall: An extensive set of rules and regulations is checked against each request made through the files firewall function. The IT department can, for example, restrict file access outside of company headquarters to select user groups. Files with specified categories can also be blocked from certain countries at the request of the user [52].

Auditing: Users and their administrators can be audited to prevent abuse and prepare for compliance audits by using Audit [52].

The secure view: The Secure View feature allows users to transmit secret information while restricting the options accessible to recipients to avoid misuse. The function eases the burden of providing certain contacts access to private material, such as virtual data rooms for merger and acquisition due diligence [30] [52].

Ransomware Protection: Through a two-step process, the Ransomware Protection App equips enterprises to combat one of the most heinous and destructive kinds of ransomware. It protects data loss by limiting sync uploads from compromised clients. Additionally, it preserves prior versions of data for rollback purposes [48] [52].

Antivirus: The ownCloud Anti-Virus App ensures that files are screened for viruses, trojans, and other dangerous code before being uploaded to the cloud.

Server-Side Security: A 4096-bit powerful private/public key pair is automatically made by the server for each person who logs in to their ownCloud account. As part of AES-256 cryptography, private keys are encrypted with a user's login password.

Each time a new file is uploaded to or synced, ownCloud produces a file-key for that file, which it then uses to encrypt the file using AES-256 encryption. As a result, every file that is known to ownCloud has a distinct file keyS. Figure 29 explains the process of encryption functions of ownCloud's server-side. Along with OwnCloud's public key, each user's public key is utilized to encrypt their file key. This encryption method generates one or more share keys. A unique share-key is provided to each file that a user may access.

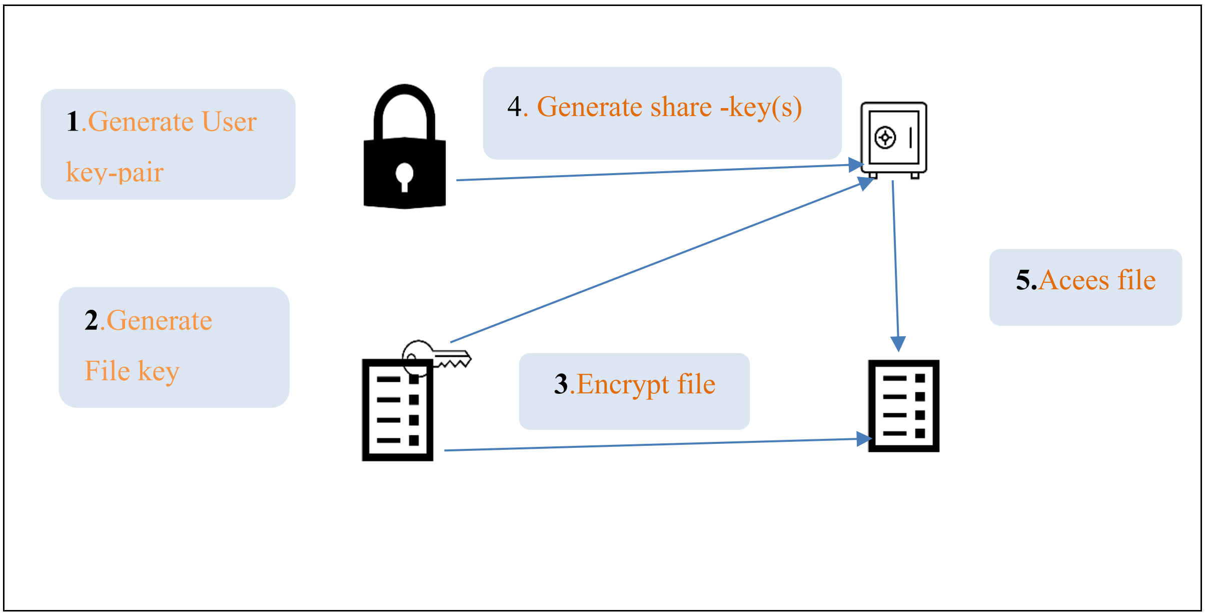

Figure 29: ownCloud's Server Side Encryption Functions [49]

An authorized user's private key and the corresponding sharing key are used to decrypt the file key when a request for access to a file is received. It is then used to unlock the encrypted data file. If a file is later shared with a different user, the file key is once again encoded using the new user's public key, resulting in the generation of a new share key. When a file is shared with new users, this abstraction forces ownCloud to re-encrypting the file key, but it avoids the far more expensive task of re-encrypting the entire physical file. When a user's access to one or more files is revoked, the same advantage is realized.

Because of its adaptive architecture and capability of handling a variety of security requirements, ownCloud's encryption paradigm ensures data security across a variety of storage formats without compromising data security that has been proven time and time again by server-side encryption. Owners of ownCloud's encryption technology have complete control over their encryption keys, which is incredibly scalable. [48], [52], [49]

3.3.2 Security and Encryption Techniques of NextCloud

The server's architecture of nextcloud is such that it has both passive and active security features. This solution implements an extensive range of security measures, including data defense strategies 'Same-site' cookies Safety precautions against bruising Web servers and OpenSSL clients use TLS encryption to encrypt and secure data flow on both sides [43].

Authentication: As part of the authentication system, which also provides pluggable validation like two-factor verification and device-oriented passkeys, a list of associated navigators (browsers) and devices are displayed on the client's private site [42]. As an added degree of security, accessibility to the storage system can be restricted to device-specific password tokens. To invalidate active sessions, remove the client from the admin panel or alter the client's credentials. Administrators can turn on or off two-factor verification for clients from the admin panel [50].SAML 2.0 (including "Shibboleth") and Kerberos verification are facilitated by Nextcloud, which also offers extensive LDAP which is an open standard tool that is applied to gain accessibility to directory information over an online environment [50].

AI Protection: NextCloud employs AI techniques to identify malicious login operations.

Rating and brute force protection: This feature restricts multiple guessing login attempts which support to protect against brute force attack. Rate limiting allows a developer to set a limit according to how many connections an IP span or a client can make in a given timeframe. This can be useful for things like reducing the cost of API queries, securing users from obtaining huge data in a short period, and hardening brute force attacks even more. Nextcloud applications safeguard members from malicious and overuse [42] [50]. Admin can impose encryption quality standards that Nextcloud will follow. When key information such as the user's email account is changed, password reset tokens are invalidated to defend against phishing attempts. Nextcloud will prompt admins for credential verification while conducting security-sensitive actions [50].

Encryption (Client-Side and Server-Side): NextCloud protects its servers with AES-256 encryption. It offers multilayered cryptographic techniques, improves data security on both sides (client and server), and exchanges data between users using TLS over HTTPS [50].

3.3.3 Security and Encryption Techniques of Seafile

In Seafile application, documents are formed into libraries in the Seafile system. Any device can be individually synced with a library. A user-defined passcode can also be applied to encode a library. Information is secure before it is transmitted to the server. The documents are also inaccessible to the admin [53].Only hashed keys are used to keep track of user credentials. When a user signs in, they must use a credential that is distinct from the credential for an encoding library.

The encryption mechanism is as follows:
1. Create a secure 32-byte Arbitrary value that will serve as the cipher key for the file ("file key").

2 Apply the client's chosen passcode to the document key and preserve the outcome. The passcode is converted into a key/iv pair using the PBKDF2 algorithm. The document key is then secured using the AES 256/CBC technique. Clients can obtain the file code from the encrypted file code if they desire accessibility to data later on.

3. The document secure code encodes the entire information of the file using AES 256/CBC cryptography. The PBKDF2 technique is utilized to acquire a key/iv pair from the document code, with the file passcode as the input. The information is transferred to the server once it has

been protected. The above-mentioned ciphering technique can be applied by client apps.[53] [51]

Each Seafile desktop application is given its private key. When a user connects to the server, they will exchange their public keys and negotiate a session key. The PDKDF2 algorithm is used to create a secure random integer for the session key. Furthermore, information is protected during transfer between the client and server by applying RSA encryption. The AES-256/CBC algorithm will be used to cipher the data transmission [51]. The goal of Seafile is to provide a dependable synchronizing tool. The synchronization technique used by Seafile is extremely robust and consistent [54].

3.4 Vulnerabilities

The need for exponential expansion in storage capacity has made storage technology one of the most important technologies. Although there are numerous advantages to using cloud storage, there are also several key obstacles to overcome. Adoption is hampered by several issues, one of which is security [55]. Vulnerabilities in the cloud have become a major cloud storage security concern. It is essential to acquire the knowledge of the risks of incorporating open-source software tools, systems, and scripts into systems before selecting any solutions. Recognizing these threats can assist users in better directing security resources and safeguarding systems. This section examines the most significant flaws in each of the three open-source cloud storage alternatives. [56]

3.4.1 Vulnerabilities of OwnCloud

The following are the significant vulnerabilities of the owncloud system.

1) Permissions, Privileges, and Access Controls: A remote intruder could gain escalated access to the system as a consequence of this flaw. This threat arises because the unified share receiver can raise capabilities, letting security limitations be avoided and privileges are elevated [57].

2) Information disclosure: A malicious actor could get accessibility to possibly sensitive information as a result of the flaw. A threat exists because of the application's massive data output. An authorized intruder with remote access can view the shareowner's internal path and username [57] [58].

3) Session Fixation: This flaw enables a local client to obtain entry into a connection to the target device. As session cookies are not reset after verifying for public connections, the vulnerability exists. A remote hacker can gain entry to the system's critical contents. [57]

4) Improper Protection of Alternate Path: A malicious hacker could get control of a compromised machine By utilizing this flaw. The API of share info can be used to avoid the privilege validation that performs when a file is dumped, which would be the vulnerability's source. An identified malicious actor can look through the files in the file drop.[57] [58]

5) OS Command Injection: Prior to ownCloud 1.0.0, the file antivirus component permitted OS Command Injection through the administration setup [57].

6) Resource Injection: In earlier version 2.9.2, a server may inject Resource Injection into the desktop client via a URL, resulting in malware [57].

3.4.2 Vulnerabilities of NextCloud

The following sections highlight the critical vulnerabilities of the Nextcloud system.

Link following: When extracting an archive, the application is not capable to examine to see if the file is a symbolic link, which generates the flaw. If the application receives a corrupted document, it can be applied to modify arbitrary information in the system. If the flaw is performed successfully, a hacker may be able to connect to the compromised machine. An intruder can connect to a vulnerable system using this flaw [59].

Authorization bypass through user-controlled key: A remote hacker could be capable to compromise the target system as an outcome of this security weakness. The flaw exists as a consequence of the inability of verification. Any Unified document share can be compromised by a remote attacker who has capable to modify it.[59] [60]

Information disclosure: this flaw exposes critical information that may be connected remotely. The lookup server could receive a client's ID even though no fields are set to be disclosed on the member's account. Using a remote attack, a hacker can get capable to connect to the computer system to access confidential information [60].

Improper access control: A remote attacker can utilize this issue to get unauthorized access to functionality that would otherwise be restricted. The flaw exists as a result of erroneous access rights measures. The admin page can be compromised by a remote authenticated attacker [60].

Improper Handling of Unexpected Data Type: This flaw could be used by a remote intruder to gain access to the target machine. The threat emerges as a consequence of the documents being dropped, and a unified share can be generated by providing a public link. Unauthorized users can modify a file drop link into a federated share, resulting in the user encountering an error when attempting to connect the shared file. [59] [60]

Code Injection due to Incorrect sanitized talk: A code injection occurs an authority incorporates an inappropriate sanitized talk instruction to Nextcloud Talk 6.0.4, 7.0.2, or 8.0.7 [60].

Vulnerable to brute force attacks: Nextcloud servers prior to 19.0.11, 20.0.10, and 21.0.2 are prone to attacks because IPv6 subnets are not regarded in rate restricting estimation. As a consequence, an intruder can avoid rate-limit protections such as Nextcloud's brute-force defense [59].

Vulnerability to webauthn tokens: Nextcloud Server is a storage solution for information. according to the versions 19.0.13, 20.011, and 21.0.3, Webauthn tokens are not discarded after a service user is erased. If the former client used an old login, they could connect to the target's profile [60].

SQL Injection: In Earlier Android app version 3.0.0 of nextcloud, SQL Injection permits a hacker to prompt to erase a local cache, attempting to enable the profile to be reset. [60]

3.4.3 Vulnerabilities of Seafile

Some significant Seafile system vulnerabilities have been described in the following sections.

Persistent XSS: The vulnerability can be exploited by cross-site scripting (XSS). The issue exists as a result of the "share of the library" functionality's insufficient data sanitization. A malicious intruder can inject, and execute HTML and malicious script in a victim's browser. An attacker can make use of this flaw to steal personal information, alter the appearance of a website, and launch phishing and drive-by-download attacks [61] [62].

Vulnerability in sync token: A sync token is used by Seafile's file synchronization mechanism to provide clients gain obtain to library information. The token is stored in memory on the seaf-server to enhance efficiency. When the server obtains a token from a sync terminal, it examines its cache to see if the token is already present. The server does not check whether the token is linked to the repository supplied in the URL if the token is present in the cache. Data from any recognized library can be accessed using any valid sync token. To begin, the attacker must find a library that it does not have access to. The library ID is a one-of-a-kind UUID that cannot be deduced from the library's URL. [62] [63]

DLL hijacking: As exchndl.dll can be fetched from the current directory, DLL hijacking is a risk with the seafile-client client 7.0.8 for Seafile [63].

Sever's vulnerability to Log4shell (log4j): The flaw in log4j takes advantage of the inadvertent processing of tampered log entries. Malicious code downloaded via the Internet can be reloaded and executed via tampered log entries. In the worst-case scenario, an intruder might abuse these security holes to acquire root access to the system and run code, such as creating a backdoor [63].

4 Result and Discussion

The study results indicate that The three open-source cloud data collaboration and synchronization systems mentioned above are powerful and flexible alternatives. All these above Dropbox rivals offer file access, synchronization, and collaboration across several devices. For a given application case, each of them is usually the best alternative [33]. The most significant findings of the thesis, their interpretation, and the implication of the study have been reviewed and assessed in this chapter. Some potential limitations and recommendations for future studies are also discussed in this chapter.

4.1 Major findings and their interpretations

The study shows a comprehensive analysis of the benefits of one service over another, as well as their drawbacks. ownCloud is designed from the ground up to be an open standard to keep up with commercial cloud providers. Nextcloud is a version of ownCloud developed by the core ownCloud team, including its founder. Seafile is a different type of file synchronization solution [64] [65]. A few technologies in Nextcloud let a team collaborate better efficiently. Collabora Online is a LibreOffice-based Office suite that is integrated with the nextcloud system to permit clients to read and update data from anywhere. The visitors' feature of ownCloud assists clients to create restricted profiles that grant visitors to interact beyond being labeled as a member, which is a unique feature. Seafile has a library feature that allows users to generate libraries of files and directories to synchronize or exchange [33].

Analyzing owncloud, nextcloud, and seafile, ownCloud comes with extensive information, including deployment and configuration instructions for users, admins, and developers, along with a GitHub repository where users may get access to the source code. The documentation website for Nextcloud includes connections to its communities, Internet Relay Chat (IRC), and social media channels for communal assistance. The community edition of Seafile is open source under the GPLv2 license. However, the professional edition is not [33]. Excluding Windows Mobile, ownCloud operates all other major platforms. Nextcloud software is compatible with Windows 7 and later, macOS 10.10 and later, Linux platforms, and mobile applications. Seafile is compatible with the same OS environment as ownCloud [33] [64].ownCloud's portfolio includes contacts applications and an official calendar. In contradiction to Nextcloud, there is no option to include text, voice, or web conference capability. Nextcloud includes technologies such as Nextcloud 'Talk' and 'Groupware' which help teams communicate and cooperate more efficiently[64]. All of these qualities are missing from Seafile. It's mostly a file synchronization and sharing platform with some online office features thrown in for good measure [65].Owncloud levies fees for collaborating with a huge number of participants. Clients who operate their servers can use OwnCloud for free. Data can be freely transmitted and synchronized between platforms. Only the server's capacity restricts the storage capacity available Anyone who manages their server can run NextCloud for free. Clients can use any of the compatible device software, as well as online apps in their preferred browser, to synchronize and link their data. NextCloud can also be operated in conjunction with third-party service providers like Dropbox [33] [64]. Seafile provides free service for up to three users, including clients hosted on their servers. Users can trade documents with others and transfer information across multiple platforms using the free edition [65].OwnCloud adds an extra layer of protection to files by encoding them with an authorization or master identifier. The end-to-end cipher feature entirely ciphers the information, thus even the admin is unable to view it. NextCloud offers better service in terms of security because it facilitates users to host their services with multilayered security mechanisms that are not commonly available with the rest [33] [64].To ensure the protection of Nextcloud servers, admins can control and implement the 2FA method on a global or group basis. End-to-end cipher techniques are also included, which are distinctive in that the user governs his or her own secure and personal audio/video communication system. Seafile ensures privacy protection, by ciphering the library with a user-defined passcode and ciphering the client end while utilizing PC terminal synchronization [33] [65].

The analysis shows that nextcloud delivers a better security solution since it enables users to host their services with comprehensive security measures that are not generally available with the rest. The majority of OwnCloud's security features are available as a free, self-hosted version. Ransomware protection is not available in the ownCloud free edition. NextCloud is needed more IT competence to build and manage the cloud solution NextCloud delivers advanced defense solutions to all users. Sensitive information is protected by end-to-end cipher, and clients have entire authority over documents and app accessibility. Furthermore, the protection of NextCloud's application is frequently being improved [64]. The seafile application employs AES 256-CBC encryption to encrypt data at rest and in transit. Aside from two-factor verification, the free edition allows users to regulate document and app access permissions [33] [65]. This thesis provides a comparison of the three open-source cloud storage options described above. The characteristics, architecture, implementation technology, security, and vulnerability concerns are all illustrations of the functional qualities of each system. The findings show that while each cloud storage solution has its own set of advantages, it is critical to understand that regardless of which option is chosen, some issues remain unsolved. The findings have demonstrated the answers to the research questions or objectives.

A study is done by IDC (International Data Corporation) in August 2008 among senior company executives and IT professionals addressing the challenges that primarily affect the success of the Cloud. And the poll results place security at the top of the list, emphasizing its significance in comparison to other cloud aspects [66]. Thus it is recommended to establish a security plan or model when implementing a self-hosted file sync and share solution, outlining precisely which hazards the server should guard against and structuring the system environment that facilitates protection against those threats. IT users should be more conscious of the numerous open-source cloud storage programs accessible, as well as how to make the best use of the technologies available to make better decisions, whether for personal or business purposes.[20] [67]

The study demonstrates a thorough examination of the advantages of one service over the other, as well as their shortcomings. This analysis supports the theory that while the three services approach confidentiality, integrity, and availability in slightly different ways, they ultimately achieve the same goal of providing their end clients with a high level of security, reliability, and efficiency. Open-source cloud storage offers flexibility, on-demand services, and a wide range of customization options. The results indicate that open-source cloud solutions offer enhanced scalability, portability, security, control, agility, and on-demand services to end-users [28] [33]. Users will be able to better understand the characteristics of the open-source cloud platforms because of the study and succinct summation and will be able to select the most appropriate services based on their needs. They will also be able to make more unified choices on the open-source cloud platform based on features, architecture, security mechanism, vulnerability, and development support. Table 3 shows an overall comparison view of three cloud storage solutions (ownCloud, Nextcloud, and Seafile) in terms of features, architecture, security, and vulnerabilities.

Table 3: Overall Comparison of ownCloud , Nextcloud, and Seafile [33], [64], [65]

Characteristics	OwnCloud	Nextcloud	Seafile
Compatible OS and devices	Linux, Windows, macOS, Android, iOS	Windows,Linux, macOS,, Windows Mobile Android, iOS	Windows, macOS, Linux, Android, iOS
Integrations & Apps	Voice & Teamware apps, App store with 120+ apps	App store with 200+ apps	Not available
Features of sharing	Distribute to a user or group, Public but secure links, Guests feature, Online workspace	Distribute a user or group, Public but secure links,Online workspace	Distribute to a user or group, Public but secure links, Library feature, Online Office
Download choices	Archive file, Distribution packages, Cloud providers,Web installer, Appliances and images,	Web installer, Distribution packages, Archive file, Appliances and images, Cloud providers, Ready-made devices	Archive file, Docker images, Distribution packages,Web installer
Deployment	On-premises, Hybrid, Cloud	On-premises, Hybrid, Cloud	On-premises, Cloud
Large file comptability	Yes	Yes	No
Features of Security	control of file access, LDAP, Active Directory, Kerberos,Storage cipher, End-to-end encryption,	End-to-end encryption,LDAP, Native SAML, Active Directory, Kerberos,File access control, App access,Storage encryption	LDAP, Shibboleth, Active Directory, Kerberos,Storage encryption,End-to-end encryption, App access rights
Client authorization through open standards (OAuth2, OpenID Connect)	The Identity Provider is used to authenticate the client	only accessible through the server	Most prominent single-sign-on authentication protocols are supported.
Storage encryption	ownCloud Enterprise is bundled with HSM support for cipher	HSM support is available on demand	Yes
Validation of file integrity and ICAP integration	Yes	No	No
Certification for storage	Scality & SUSE	No	Yes
Document sorting	Yes	No	Yes
Assurance of Quality	Devoted QA team	Engineering is ingrained	Debian Seafile Team
Kanban Support	No	Yes	No
Scripted Automation	Yes	Yes	No
Real Time Notifications	No	Yes	No
Multi-factor authentication	Yes	Yes	No

4.2 Implications of the study

The findings reveal some significant distinctions in cloud storage platform preferences. As a result, the findings can be used as guidelines when choosing an open-source cloud storage solution. This thesis contributes to a clearer understanding of these three open-source cloud storage solutions and their strengths and limitations so that end-users can choose their open-source cloud storage solution according to their demands. This research is a significant step toward defining open and adaptable benchmarking approaches for open source cloud storage solutions. It can assist end-users in comparing options and making knowledgeable choices based on features, architecture, security, and vulnerabilities. Finally, the findings also provide comparison tools to contribute to the research community to encourage additional research into open source cloud storage architecture, security, and performance.

4.3 Limitations and Future Research

This study may have some potential limitations. One of the most noticeable limitations is the absence of survey data. A survey of professionals and end-users could provide additional relevant information regarding these three cloud solutions. As a result, the survey data may improve the quality of study findings. Surveys would provide a credible dataset from which to derive specific results and inferences crucial decisions. Thus this issue should be addressed in future studies to improve the quality of the results. Further study can be conducted to evaluate other prominent open-source cloud storage systems using various methodologies such as systematic literature reviews and surveys of open source cloud storage specialists and end-users.

5 Conclusion

Cloud storage is a more contemporary internet-based technology that has been widely used and investigated. Data security and privacy issues have become increasingly important as the volume of data stored on cloud servers has grown. Cloud users must assess the security of cloud storage solutions. Apart from commercial services, there are numerous open-source cloud storage options available, each with a different level of application, architecture, features, security, and vulnerabilities. The distinctions among today's open-source options are only in the details: different features, different administration tools, and different security techniques [28].

Although there are numerous open-source cloud storage services for industrial and academic applications, it is difficult for individuals or an organization and their users to select the most appropriate platform for their needs. A detailed evaluation of three main open-source cloud storage options is offered in this thesis [68]. The comparison was made based on the current features, architecture, security techniques, vulnerabilities, and technologies of open-source cloud storage software. The platform's differences might be confusing now in terms of comprehension and usage. Readers may now comprehend the characteristics and make cloud storage solution selections based on the analysis in terms of cloud storage modules, services, development support, cloud interfaces, deployment, OS support, security, vulnerabilities, and compatibility.

Even though each cloud storage option has its own set of advantages, It is important to remember that no matter whether the option is chosen, several issues have yet to be resolved. Maintaining high reliability, dealing with cluster failure approaches in a cloud system, assuring continuity, synchronization across various clusters in cloud infrastructure, interoperability, standards, and cloud system protection are all areas that need improvement [67][68]. Those issues can be addressed in further studies. Further research is needed on these concerns to enhance efficiency and mitigate security problems in cloud storage for future technology.

References

[1] N. Akhtar, B. Kerim, Y. Perwej, A. Tiwari, and S. Praveen, ‘A Comprehensive Overview of Privacy and Data Security for Cloud Storage’, *Int. J. Sci. Res. Sci. Eng. Technol.*, 2021.

[2] A. Oussama and Z. Abdelhafid, ‘Cloud Storage and Security Overview’, 2018, Accessed: Dec. 20, 2021. [Online]. Available: http://ceur-ws.org/Vol-2326/paper3.pdf

[3] M. Gregus and V. Karovic, ‘Practical Implementation of Private Cloud Based on Open Source ownCloud for Small Teams - Case Study’, in *2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, Krakow, Poland, Nov. 2015, pp. 183–187. doi: 10.1109/3PGCIC.2015.149.

[4] E. Torres, G. Callou, and E. Andrade, ‘A hierarchical approach for availability and performance analysis of private cloud storage services’, *Computing*, vol. 100, no. 6, pp. 621–644, 2018.

[5] H. Snyder, ‘Literature review as a research methodology: An overview and guidelines’, *J. Bus. Res.*, vol. 104, pp. 333–339, Nov. 2019, doi: 10.1016/j.jbusres.2019.07.039.

[6] Auraria Library, ‘Research Methods: Literature Reviews’. Accessed: May 02, 2022. [Online]. Available: https://guides.auraria.edu/researchmethods/literaturereviews

[7] J. Wu, L. Ping, X. Ge, Y. Wang, and J. Fu, ‘Cloud Storage as the Infrastructure of Cloud Computing’, in *2010 International Conference on Intelligent Computing and Cognitive Informatics*, Kuala Lumpur, Malaysia, Jun. 2010, pp. 380–383. doi: 10.1109/ICICCI.2010.119.

[8] T. KamalaKannan, K. Sharmila, M. C. Shanthi, and M. R. Devi, ‘Study on Cloud Storage and its Issues in Cloud Computing’, *Int. J. Manag. Technol. Eng.*, vol. 9, no. 1, pp. 976–981, 2019.

[9] R. A. P. Rajan, ‘Evolution of Cloud Storage as Cloud Computing Infrastructure Service’, *IOSR J. Comput. Eng.*, vol. 1, no. 1, pp. 38–45, 2012, doi: 10.9790/0661-0113845.

[10] R. Nachiappan, B. Javadi, R. N. Calheiros, and K. M. Matawie, ‘Cloud storage reliability for Big Data applications: A state of the art survey’, *J. Netw. Comput. Appl.*, vol. 97, pp. 35–47, Nov. 2017, doi: 10.1016/j.jnca.2017.08.011.

[11] Dr. N. Akhtar, Dr. B. Kerim, Dr. Y. Perwej, Dr. A. Tiwari, and Dr. S. Praveen, ‘A Comprehensive Overview of Privacy and Data Security for Cloud Storage’, *Int. J. Sci. Res. Sci. Eng. Technol.*, pp. 113–152, Sep. 2021, doi: 10.32628/IJSRSET21852.

[12] D. V. Bhavsagar, V. Chavan, and S. J. Sharma, ‘Evolution of Multi Cloud Framework for Integrity, Confidentiality and Availability’.

[13] I. Ashraf, ‘An overview of service models of cloud computing’, *Int. J. Multidiscip. Curr. Res.*, vol. 2, no. 1, pp. 779–783, 2014.

[14] P. S. Suryateja, ‘Threats and vulnerabilities of cloud computing: a review’, *Int. J. Comput. Sci. Eng.*, vol. 6, no. 3, pp. 297–302, 2018.

[15] J. Gibson, R. Rondeau, D. Eveleigh, and Q. Tan, ‘Benefits and challenges of three cloud computing service models’, in *2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)*, 2012, pp. 198–205.

[16] M. U. Bokhari, Q. M. Shallal, and Y. K. Tamandani, ‘Cloud computing service models: A comparative study’, in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 890–895.

[17] R. Ara, M. A. Rahim, S. Roy, and U. K. Prodhan, ‘Cloud Computing: Architecture, Services, Deployment Models, Storage, Benefits and Challenges’, *Int. J. Trend Sci. Res. Dev.*, vol. 484, pp. 837–842, 2020.

[18] L. Savu, ‘Cloud Computing: Deployment Models, Delivery Models, Risks and Research Challenges’, in *2011 International Conference on Computer and Management (CAMAN)*, Wuhan, China, May 2011, pp. 1–4. doi: 10.1109/CAMAN.2011.5778816.

[19] Assistant Professor, Achraya Motibhai Patel Institute of Computer Studies Ganpat University, Kherva India, Prof. H. B. Patel, Prof. N. Kansara, and Assistant Professor, JG College of Computer Applications, Gujarat University, India, ‘Cloud Computing Deployment Models: A Comparative Study’, *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 9, no. 2, pp. 45–50, Mar. 2021, doi: 10.21276/ijircst.2021.9.2.8.

[20] A. Olaosebikan, ‘Security & Privacy Comparison of NextCloud vs Dropbox: A Survey’, p. 7.

[21] W. Zeng, Y. Zhao, K. Ou, and W. Song, ‘Research on cloud storage architecture and key technologies’, in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, 2009, pp. 1044–1048.

[22] G. Kulkarni, R. Waghmare, R. Palwe, V. Waykule, H. Bankar, and K. Koli, ‘Cloud storage architecture’, in *2012 7th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*, Denpasar-Bali, Indonesia, Oct. 2012, pp. 76–81. doi: 10.1109/TSSA.2012.6366026.

[23] F. Yahya, R. J. Walters, and G. B. Wills, ‘Using goal-question-metric (GQM) approach to assess security in cloud storage’, in *International Workshop on Enterprise Security*, 2015, pp. 223–240.

[24] M. Kaur and H. Singh, ‘A review of cloud computing security issues’, *Int. J. Adv. Eng. Technol.*, vol. 8, no. 3, p. 397, 2015.

[25] V. Pai T. and P. S. Aithal, ‘Cloud Computing Security Issues - Challenges and Opportunities’, *Int. J. Manag. Technol. Soc. Sci.*, pp. 33–42, Dec. 2016, doi: 10.47992/IJMTS.2581.6012.0004.

[26] F. Yahya, V. Chang, R. J. Walters, and G. B. Wills, ‘Security Challenges in Cloud Storages’, in *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, Singapore, Singapore, Dec. 2014, pp. 1051–1056. doi: 10.1109/CloudCom.2014.171.

[27] N. Subramanian and A. Jeyaraj, ‘Recent security challenges in cloud computing’, *Comput. Electr. Eng.*, vol. 71, pp. 28–42, 2018.

[28] M. K R and R. swamy, ‘Cloud Computing Applications: An Open Source Software (OSS) Approach’, *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3507358.

[29] I. Syamsuddin, R. Nur, D. Nur, Irmawati, and D. Al-Dabass, ‘LOW COST e-GOVERNMENT CLOUD DATA CENTER: A CASE STUDY OF OPEN SOURCE ADOPTION’, *Far East J. Electron. Commun.*, vol. 18, no. 2, pp. 237–250, Mar. 2018, doi: 10.17654/EC018020237.

[30] ownCloud, ‘All Features’. Accessed: Feb. 05, 2022. [Online]. Available: https://owncloud.com/features/

[31] Nextcloud, ‘Key capabilities’. Accessed: Feb. 12, 2022. [Online]. Available: https://nextcloud.com/hub/

[32] Seafile, ‘Seafile Features’. Accessed: Feb. 14, 2022. [Online]. Available: https://www.seafile.com/en/features/

[33] Nextcloud, ‘Compare products’. Accessed: Apr. 06, 2022. [Online]. Available: https://nextcloud.com/compare/

[34] ownCloud, ‘ownCloud Architecture Overview’. Accessed: Feb. 07, 2022. [Online]. Available: https://owncloud.com/wp-content/uploads/2014/03/oc_architecture_overview.pdf

[35] P. Gopalakrishnan and B. U. Maheswari, ‘Research on enterprise public and private cloud service’, *Int J Innov Technol Explore Eng*, vol. 8, no. 6, pp. 1453–1459, 2019.

[36] I. Syamsuddin, A. Satria Prabuwono, A. Hoirul Basori, and A. Yunianta, ‘Review on OwnCloud Features for Private Cloud Data Center’, *TEM J.*, pp. 954–960, May 2021, doi: 10.18421/TEM102-59.

[37] N. Singh, K. Bui, and A. Mailewa, ‘Robust Efficiency Evaluation of NextCloud and GoogleCloud’, *Adv. Technol.*, pp. 536–545, 2021.

[38] S. Asenov, V. Raydovska, S. Lyubomirov, and D. Shehova, ‘USING OF THE NEXTCLOUD TECHNOLOGY IN THE ENGINEER EDUCATION’, Seville, Spain, Nov. 2019, pp. 7087–7095. doi: 10.21125/iceri.2019.1684.

[39] S. R. Siregar and P. Pristiwanto, ‘Build Data Backup with Nextcloud Based Infrastucture as A Service (IAAS) Concept on Budi Darma University’, *IJICS Int. J. Inform. Comput. Sci.*, vol. 5, no. 1, pp. 95–101, 2021.

[40] H. Terauchi and N. Xiong, ‘An Effective Seafile Dockerfile for Raspberry Pi to Make Docker YAML Files for Treehouses’, in *Smart Computing and Communication*, vol. 12608, M. Qiu, Ed. Cham: Springer International Publishing, 2021, pp. 127–135. doi: 10.1007/978-3-030-74717-6_14.

[41] P. Stąpór, D. Laskowski, and P. Łubkowski, ‘Private Cloud Architecture - Analysis of Reliability’, *J. KONBiN*, vol. 45, no. 1, pp. 267–286, Mar. 2018, doi: 10.2478/jok-2018-0014.

[42] Nextcloud, ‘History and Architecture’. Accessed: Feb. 20, 2022. [Online]. Available: https://docs.nextcloud.com/desktop/3.4/architecture.html

[43] Nextcloud, ‘Nextcloud Solution Architecture’, Accessed: Feb. 28, 2022. [Online]. Available: https://nextcloud.com/media/wp135098u/Architecture-Whitepaper-WebVersion-072018.pdf

[44] Doğan Can Uçar, ‘Design and Implementation of a File Recommendation System Using Collaborative Filtering and Content-Based Recommendation for the Nextcloud Platform’, 2018, doi: 10.13140/RG.2.2.14620.31360.

[45] Seafile, ‘Deploy in a cluster’. Accessed: Mar. 10, 2022. [Online]. Available: https://manual.seafile.com/deploy_pro/deploy_in_a_cluster/

[46] Jonathan Xu, ‘Seafile 2020 & Future Development’. Accessed: Mar. 15, 2022. [Online]. Available: https://indico.cern.ch/event/970232/contributions/4157916/attachments/2176958/3676224/seafile-cs3-2021-talk.pdf

[47] S. Yang, L. Jiang, S. Zhu, and L. Dai, ‘Research and application of private cloud storage platform in high schools based on seafile’, in *2013 6th International Conference on Intelligent Networks and Intelligent Systems (ICINIS)*, 2013, pp. 25–28.

[48] B. Rexha, B. Likaj, and H. Lajqi, ‘Assuring security in private clouds using ownCloud’, *Ijacit Com*, 2012.

[49] ownCloud, ‘ownCloud’s Data Encryption Model’. Accessed: Mar. 20, 2022. [Online]. Available: https://owncloud.com/wp-content/uploads/2014/10/Overview_of_ownCloud_Encryption_Model_1.1.pdf

[50] J. Poortvliet, ‘Encryption in Nextcloud’, Feb. 05, 2018. https://nextcloud.com/blog/encryption-in-nextcloud/ (accessed Mar. 21, 2022).

[51] Seafile, *Security Features*. Accessed: Mar. 24, 2022. [Online]. Available: https://lins05.gitbooks.io/seafile-docs/content/security/security_features.html

[52] ownCloud, ‘Ensure Security’. Accessed: Mar. 17, 2022. [Online]. Available: https://owncloud.com/product/security/

[53] Seafile, ‘Encrypted Library’. Accessed: Mar. 23, 2022. [Online]. Available: https://manual.seafile.com/security/security_features/

[54] M. Dreyer and A. BrinNmann, ‘Building an Online File Storage and Sharing Service using Seafile and cEPH for HU Berlin and JGU Mainz’. Berlin, 2015.

[55] B. Grobauer, T. Walloschek, and E. Stocker, ‘Understanding cloud computing vulnerabilities’, *IEEE Secur. Priv.*, vol. 9, no. 2, pp. 50–57, 2010.

[56] M. Derfouf, A. Mimouni, and M. Eleuldj, ‘Vulnerabilities and storage security in cloud computing’, in *2015 International Conference on Cloud Technologies and Applications (CloudTech)*, 2015, pp. 1–5.

[57] Ö. Serkan, ‘Owncloud : Security Vulnerabilities’. Accessed: Mar. 26, 2022. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-11929/Owncloud.html

[58] W. Xu, B. Groves, and W. Kwok, ‘Penetration testing on cloud---case study with owncloud’, *Glob. J. Inf. Technol.*, vol. 5, no. 2, p. 87, Jan. 2016, doi: 10.18844/gjit.v5i2.198.

[59] Ö. Serkan, ‘Nextcloud : Security Vulnerabilities’. Accessed: Mar. 28, 2022. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-15913/product_id-34622/Nextcloud-Nextcloud.html

[60] stach.watch, ‘Nextcloud’. Accessed: Mar. 30, 2022. [Online]. Available: https://stack.watch/product/nextcloud/nextcloud/

[61] stack.watch, *Seafile*. Accessed: Apr. 01, 2022. [Online]. Available: https://stack.watch/product/seafile/seafile/

[62] Ö. Serkan, *Seafile : Security Vulnerabilities*. Accessed: Apr. 01, 2022. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-21226/product_id-64019/Seafile-Seafile.html

[63] Vulmon, ‘seafile vulnerabilities and exploits’. Accessed: Apr. 03, 2022. [Online]. Available: https://vulmon.com/searchpage?q=seafile&sortby=byrelevance

[64] ownCloud, ‘ownCloud vs Nextcloud’. Accessed: Apr. 05, 2022. [Online]. Available: https://owncloud.com/owncloud-vs-nextcloud/

[65] ownCloud, ‘Seafile vs ownCloud’. Accessed: Apr. 05, 2022. [Online]. Available: https://owncloud.com/seafile-vs-owncloud/#:~:text=ownCloud%20has%20a%20large%2C%20global,problem%20has%20already%20been%20adressed.

[66] P. Deepanchakaravarthi and Dr.Sunitha Abburu, ‘An Approach for Data Storage Security in Cloud Computing’, Accessed: Apr. 12, 2022. [Online]. Available: https://www.researchgate.net/publication/265973909_An_Approach_for_Data_Storage_Security_in_Cloud_Computing

[67] ‘A REVIEW ON CLOUD STORAGE SECURITY’, Accessed: Apr. 12, 2022. [Online]. Available: https://www.researchgate.net/publication/326234875_A_REVIEW_ON_CLOUD_STORAGE_SECURITY

[68] D. Zhe, W. Qinghong, S. Naizheng, and Z. Yuhan, ‘Study on Data Security Policy Based on Cloud Storage’, in *2017 IEEE 3rd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS)*, Beijing, China, May 2017, pp. 145–149. doi: 10.1109/BigDataSecurity.2017.12.