CompressoGalaxy: The Ultimate Cloud Evolution

Prof. Swati Paralkar[1], Atharva Ajay Kale[2], Janhavi Sawarkar[3], Vivek Balasaheb Auti[4], Harshada Vijay Jagtap[5]

JSPM's RSCOE
Pune, Maharashtra, India.

Abstract:
Cloud computing is like a revolution in the software industry. However, the revolution is incomplete because people don't use cloud storage efficiently. Compresso is a cloud-based application that provides cloud services. As we all know, the cloud is an inevitable part of life; to simplify this, Compresso has evolved. Cloud Computing is highly available with storage time. It is reliable, scalable, agile, and environmentally sustainable. Compresso is an application that stores files, photos, videos, and documents in a compressed form, and if a user wants the original form of their data, they can access it at any time. Compression of Data reduces the size of the data. Compression is a technique that converts the actual data into a compressed form. If the data is stored in compressed form, then the users can use the reduced space. Everyone knows that a huge amount of energy is required in this process, and due to this cost is also high. Here, individuals require so many data center that needs to be coupled with the system, which is the main reason for the high consumption of energy. Hence, the application introduced green cloud computing, in which it reduces the energy consumption by the system, and new technologies and algorithms are used. Users will move data and application software to the cloud, hence ensuring integrity and privacy with the help of the Security Onion method. Users should trust the system, so it needs to secure the data that users will store in the system; hence, Onion Security is used. Security Onion is a process that secures the system with the help of multiple intrusion Detection Systems. Onion Security is used to secure the data that is stored. Hence, this brief talk will outline concerns about the application and the overview of the system.

Keywords: Cloud computing, Compression, Onion Security, De-Duplication, OpenCV, Pillow, Encryption-Decryption.

1. INTRODUCTION

Cloud computing plays a critical role in modern software systems and is widely integrated into everyday activities. It provides users with on-demand storage solutions, offered by major companies such as Google, Amazon, and Microsoft, through subscription-based models. Instead of relying on physical hardware, users can store data online via the cloud, enabling better resource utilization, a key advantage of cloud computing. Data compression is a feature that optimizes storage by reducing file sizes, allowing more data to be stored. However, Data Quality is affected due to compression, which results into a challenge. Despite this, our system delivers the original and unaltered data to users whenever they access it. Green cloud computing utilizes energy-efficient algorithms that minimize power consumption. Leading companies like
Amazon and Google, for various file types, including images, videos, audio, and documents, have developed services and provide compression options. While compressed data doesn't have quality, these techniques have changed how users store and manage data. Platforms such as AWS Elastic MapReduce (EMR), Google Big Query, and Azure HDInsight provide data processing techniques and also analyse the stored data, while hybrid clouds of AWS and Azure offer flexibility. Cloud computing allows users to control storage and servers with ease. Cloud computing uses software rather than spending money on physical hardware. This system ensures that functions are performed efficiently while maintaining security, prioritizing reliability. Ultimately, cloud computing enables users to store and access data anytime, anywhere, using the internet.

1.1 Cloud deployment models

Cloud computing involves using web-based systems to access a huge pool of resources. Users obtain these services over a network, whether through the internet or an internal network (intranet). Essentially, it allows individuals or businesses to utilize or purchase computing services from third-party providers. These resources can include any IT-related tools such as applications, computing storage, databases, and advanced technologies. With a huge pool of resources, users have the flexibility to scale their applications to any size, maintaining full control over their usage.

1.2 Types of cloud resources

There are four types of cloud:

1. **Public cloud:** The cloud whose resources can be shared with anyone and anyone can use it. Example: Google Aps, SQL Azure, etc.
2. **Private cloud:** This cloud resources are shared with limited people and those people can only use it.
3. **Hybrid cloud**: This is a mixture of public and private cloud. Some resources of this cloud are shared outside, some are shared with limited people.
4. **Community cloud**: This cloud computing resource is shared with more than one community that reduces the cost and makes it more efficient for organizations using this cloud.

2. OBJECTIVES:

The main objectives of this project are as follows:

1. **To develop a secure and efficient cloud storage platform** that allows users to upload, compress, and store multimedia content such as photos and videos.
2. **To design a user-friendly web interface** using HTML, CSS, and JavaScript for seamless file upload and retrieval by users.
3. **To implement an advanced media compression mechanism** using Python libraries (Pillow and OpenCV) to reduce file sizes without significant quality loss.
4. **To host the storage system on a private cloud** running on a Linux-based operating system, ensuring more control over data management.
5. **To integrate the Onion Security System (OSS)** for user credentials and sensitive data, using multi-layered encryption techniques for enhanced security.
6. **To optimize cloud resource usage** by reducing the storage footprint through effective compression while maintaining accessibility and integrity of the files.

3. ANALYSIS OF PROBLEM

The requirement analysis for this project outlines the necessary functional, non-functional, and system requirements for the cloud storage platform. Key functionalities include user authentication, media file compression using Pillow and OpenCV, and secure file retrieval. Non-functional requirements focus on performance, security, and scalability, with an emphasis on fast processing, strong encryption, and responsive design. The platform will operate on an Oracle VM with Ubuntu 64-bit OS, using Python, FastAPI, and MySQL for smooth user experience and efficient data management.

Duplication image: In this process, multiple images are duplicate. We use de duplication algorithm for duplicate image.

4. LITERATURE REVIEW:

In 2023, This Research focuses on different Image Compression methods to efficiently use the storage in cloud computing environments. The authors specified different image compression techniques like Discrete Cosine Transform, Discrete Wavelet Transform, Singular Value Decomposition, and Karhunen-Loève Transform. Cloud Storage usage increases, but image Quality is not ensured [1].

In 2021, Mohd Tajammul, Rafat Parveen, Nitin K Gaur, and Saravanan D discussed integrating data-sensitive algorithms with compression techniques for efficient cloud storage utilization and ensuring data security. The techniques aim to tackle the dual challenge of ensuring data security while optimizing cloud storage. It uses

Encryption Techniques that is Advanced Encryption Standard and a compression technique that is Discrete Cosine Transform [2].

The Authors proposed a Research Article on "Image Processing with Python Libraries," which uses various Python libraries utilized for image processing, with their functionalities and applications. It demonstrates how libraries like OpenCV facilitate complex operations such as image recognition and video analysis [3].

Yasha Jyothi M Shirur, Nithin Iyer K S, Sujay K S, and Uday V N proposed an article addressing the increasing need for Data Security mechanisms in today's world. Traditional Single Layer methods may not provide sufficient protection so they developed more advanced cryptographic techniques like Elliptic Curve Cryptography [4].

Jinan Shen, Xuejian Deng, and Zhenwu Xu proposed a Research Article on "Multi-security-level cloud storage system based on improved proxy re-encryption" that integrates AES symmetric encryption with an enhanced identity-based proxy re-encryption algorithm to ensure Data Security [5].

In 2022, This Research aims to strengthen the efficiency and effectiveness of image processing tasks by utilizing OpenCV's capabilities, particularly in handling large data in big data applications. OpenCV with deep learning algorithms is addressed to optimize image processing [6].

In 2023, Alireza Sajadinia and Alireza Yari proposed a Research Article that tackles the issues of upgrading virtual machine (VM) placement in cloud computing environments to enhance performance. It also includes various approaches applied to Virtual Machine placements [7].

Ku. Swati G. Anantwar and Prof. Karuna G. Bagde proposed an article that addresses the concerns and problems related to ensuring Data integrity and security in cloud computing environments. It improves security by encrypting hash values. It also helps customer to check the correctness of their data [8].

5. LIBRARIES USED:

a. OpenCV:

OpenCV (Open-Source Computer Vision Library) is an open-source library widely used for image processing tasks. It provides various tools for manipulating images and videos, which include functionalities like image compression and color space transformations. In this system, this Python library is used to improve the efficiency of image compression, allowing good image quality along with size reduction. It helps to enable the system to handle complex image operations quickly. OpenCV's performance makes it an important component in achieving the goal of efficient cloud storage management.

b. Pillow:

Pillow is a widely used Python library for image processing. It is used to handle image manipulation efficiently. It provides functionality for editing. It saves images in multiple formats such as JPEG, PNG, BMP, and GIF. This library supports operations like resizing, rotating, and filtering images, required for image optimization and image compression. In this project, Pillow plays a vital role in reducing the file size of uploaded images while maintaining quality. Its versatile nature helps in handling complex image processing for contributing to the project's goal of maximizing storage space.

6. COMPARATIVE STUDY

Data Compression	Google Drive	Dropbox	OneDrive	iCloud	Proposed Platform
Type of Cloud	Public	Public	Public	Public	Private Cloud
Data	Automatic	No	Limited	No	Advanced

Compression					(OpenCV + Pillow)
Security	TLS, AES 128/256-bit encryption	SSL/TLS, AES 256-bit Encryption	SSL/TLS, AES Encryption	SSL/TLS, AES Encryption	Onion Security System (Multi-layer Cipher)
Customizable Backend	No	No	No	No	Fully Customizable
Programming Stack	Proprietary	Proprietary	Proprietary	Proprietary	HTML, CSS, JS, Python (Pillow, OpenCV), Linux
Cost	on a subscription basis	on a subscription basis	on a subscription basis	on a subscription basis	Self-hosted (No monthly fees)
File Type Support	All formats	All formats	All formats	Apple preferred format + general formats	Photos & Videos (Customizable)
Compression Control	No	No	No	No	User-defined/ Automatic
Data Ownership	Stored on third-party servers	Stored on third-party servers	Stored on third-party servers	Stored on third-party servers	User controls private cloud

The Proposed Cloud Platform provides better features compared to other cloud storage services. Unlike these public cloud stores data on third-party servers, the proposed System allows users to control their data. It includes image compression techniques like Pillow and OpenCV, which help reduce the file size of various file types by 35-50%. For security, it uses a multi-layer Onion Security in which double encryption is used so that the data remains secure. This Proposed System makes it 40-60% more secure than other encryption used in public cloud services.

Another benefit is flexibility, reliability, and cost savings. In this system, there is no fees or subscription required, while public cloud services require monthly payments. This makes it 100% cost-efficient. With better functionalities, security, efficient cloud storage, and cost efficiency, this platform is an authentic alternative to public cloud storage.

7. SOFTWARE REQUIREMENTS:

1. Operating System: Ubuntu 64-bit (running on Oracle VM)
2. Virtualization Software: Oracle VirtualBox or VMware
3. Programming Language: Python 3.x
4. Web Framework: FastAPI
5. Database: MySQL (or Firebase for optional cloud database integration)

6. Image Processing Libraries: Pillow, OpenCV
7. Encryption Libraries: cryptography or hashlib
8. Compression Utilities: zipfile, shutil
9. Web Server: Uvicorn (for running FastAPI)
10. Version Control: Git/GitHub for code management
11. Browser: Chrome, Firefox, or any modern browser for testing

8. HARDWARE REQUIREMENTS:

1. Processor: Intel Core i5 or higher
2. RAM: 8 GB minimum (16 GB recommended for higher performance)
3. Storage:
   100 GB for the host machine
   20 GB allocated for the Oracle VM
4. Network: Stable internet connection for database and server communication
5. Display: 1080p resolution or higher for UI testing
6. Peripherals: Standard keyboard and mouse

9. PROPOSED SYSTEM

The proposed system provides a solution for the efficient management of cloud storage, combining compression and security. It manages the user's data by providing them a secure experience, and the platform is designed to ensure that it meets future demands. With a strong analysis of the performance and security, the system helps to deliver a modern and efficient solution for cloud storage.

10. USED METHODOLOGY

a. Onion Security

Onion Security System is a multi-layered security mechanism that ensures the user's data is encrypted multiple times, preventing it from various cyber attacks. Each layer is secured by double encryption techniques, protecting the data from unauthorized access. This results in making the system secure and efficient cloud storage platform.

b. Image Compression

We use image compression techniques using Python libraries like OpenCV and Pillow that help to reduce the size of image files stored in the cloud. It saves storage space. These techniques are used to preserve essential image quality and save storage space.

c. De-Duplication and Uploading

De-duplication removes unnecessary copies of data stored in the cloud. By using techniques like the SHA Hash Algorithm, the same files are automatically identified and the one with the better quality is kept, and the other is removed with the prior permission of the user, hence significantly reducing the storage and enhancing the storage efficiency.

d. Compression of Files

This includes reducing the file size through lossless or lossy compression algorithms. Compression helps to optimize storage usage while maintaining data accessibility. Compression is done using Python libraries like OpenCV and Pillow that decrease the file size.

e. Green Cloud Computing

Green cloud computing focuses on reducing energy consumption and operational costs, which results in decreasing the environmental impact. There are various algorithms used to minimize the power consumption, therefore giving a low operational cost. This is an approach that makes the system cost-efficient and environmentally friendly.

f. Effective Storage Use Techniques

These methods help to increase the efficiency of storage systems:

- Duplicate image deletion: In this system, there is a concept of automatically identifying duplicate data and removing it with the permission of the user.
- Recently deleted data removal: When the user removes some files and stores them in the recycle bin, those files are permanently deleted after a set period, which frees up space.

- Frequently used data management: Here, we archive older and least used files and move frequently used files to a folder that helps user to easily access and also use space efficiently.

g. Data Encryption

Encryption ensures the integrity of the user's data. Encryption algorithms protect files from unauthorized access. In this system, we use a double encryption mechanism called "Onion Security Model". By combining these techniques, the Platform provides a secure and efficient storage for various file types like images, videos, audio, and documents.

11. SYSTEM ARCHITECTURE

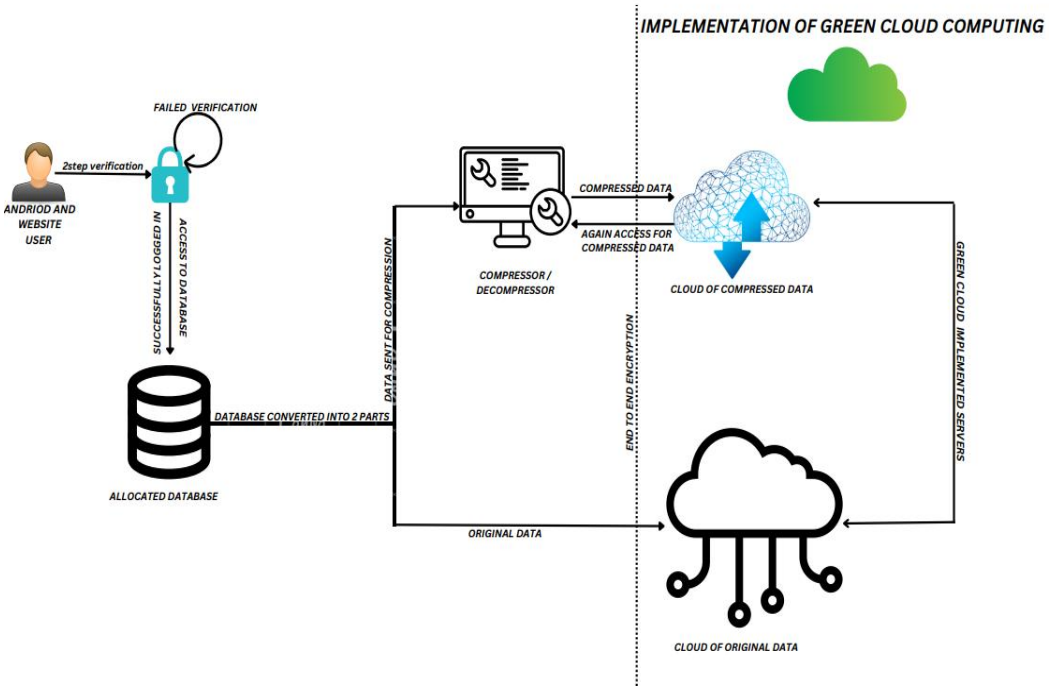

Fig. 1. System Architecture of the system

12. IMPLEMENTATION DEMONSTRATION

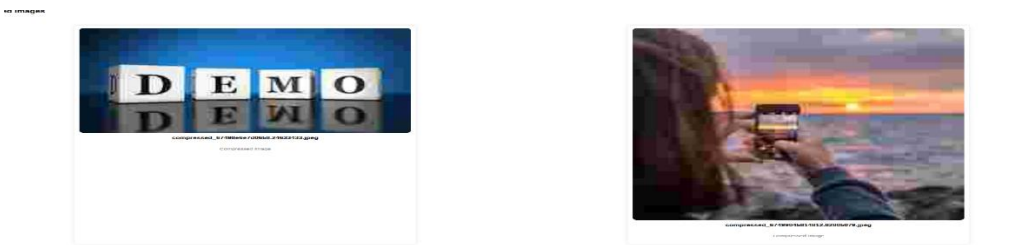

Fig. 2. Image Compression

13. CONCLUSION

This project focuses on allowing users to store, and access compressed images and videos that helps to develop a secure and efficient cloud storage. It uses image processing techniques using Python libraries such as Pillow and OpenCV to reduce file sizes without compromising quality. To ensure the security of user data, the platform uses double-encryption mechanism called the "Onion Security System" for securing the data.

This system securely stores both the original and compressed user data, therefore resulting in an efficient and cost-effective storage management.

REFERENCES:

1. Tri Yulianto Nugroho, Ahmad Nur Hidayat, Bagus Jati Santoso, "A Comparative Study on Image Compression in Cloud Computing" 31 October 2023.
2. Mohd Tajammul, Rafat Parveen, Nitin K Gaur, Saravanan D, "Data Sensitive Algorithm Integrated with Compression Technique for Secured and Efficient Utilization of Cloud Storage" 02 November 2021.
3. Renas Rajab Asaad, Rasan Ismael Ali, Zeravan Arif Ali, Awaz Ahmad Shaban "Image Processing with Python Libraries" June 2023.
4. Yasha Jyothi M Shirur, Nithin Iyer K S, Sujay K S, Uday V N, "Design and Implementation of Synthesizable Two-Level Cryptosystem for High-Security enabled Applications" 10 April 2023.
5. Jinan Shen, Xuejian Deng, Zhenwu Xu, "Multi-security-level cloud storage system based on improved proxy re-encryption" 18 December 2019.
6. WeijUN Wang, Junquan Wang, Zhiqiang Zhang, Da Shi, "OpenCV Implementation of Image Processing Optimization" 27 April 2022.
7. Alireza Sajadinia, Alireza Yari "Virtual Machine Placement Strategy Using Clustering and Genetic Algorithm for increasing cloud performance and power saving" 26 April 2023.
8. Ku.Swati G. Anantwar, Prof. Karuna G. Bagde, "Data Integrity and Security in Cloud" 7 September 2012.