

HAEDGE AUTOMATED DATABASE FAILOVER IN ENTERPRISE DATACENTRES

Sujith M¹, Siddiq kesavan S², Guru prasadh M³

¹UG Student, Department of Electrical and Electronics Engineering,

^{2,3}UG Student, Department of Electronics and Communication Engineering,

^{1,2,3}Bannari Amman Institute of Technology

Abstract

Globally distributed cloud data centers are the primary means of achieving data storage and service delivery. Nonetheless, data centers are vulnerable to failure threats brought on by calamities like disc failure. Ensuring fault tolerance is crucial for cloud computing systems to ensure dependability and availability. Additionally, data backup and recovery are critical concerns, with the demand for effective data recovery approaches growing daily. Users are starting to choose cloud backup as their method of supporting catastrophe recovery. Thus, cloud data centers that are disaster-resistant are anticipated. Using early warning time, an emergency protection plan that combines data backup and service transfer is suggested. VirtualBox is used in data recovery approaches to recover data from the HAEdge when the server is unable to supply it to users or if it has been lost due to various types of problems. This project combined HAEdge with VirtualBox local storage devices to put up a flexible data backup system. The findings demonstrate that enabling users to access backups from any platform or device with web browsing capabilities can help accomplish the goal of an ideal recovery time. Additionally, each client may achieve a high level of integrity, reducing the likelihood of data loss or financial record exposure to an attacker.

Keyword: Cloud, DataCenter, DataBase, HAEdge, VirtualBox, Backup System.

1. INTRODUCTION

One of the most important resources that a business may have is data. The cloud is among the greatest places to keep these resources. Cloud computing is the pay-as-you-go, on-demand delivery of IT services over the Internet. Instead of buying, running, and maintaining physical data centres and servers, you may get technical services like databases, processing power, and storage from a cloud provider like Amazon Web Services (AWS) on an as-needed basis.

1.1 Cloud Service Types

Cloud computing is not a singular technical apparatus, like a microprocessor or smartphone. As opposed to this, the system

is primarily composed of three services: software-as-a-service (SaaS), infrastructure-as-a-service (IaaS), and platform-as-a-service (PaaS).

1.1.1 SaaS (software as a service)

It deals with granting users of a software programme a license. Usually, Pay-as-you-go or as-needed licences are available. This kind of technology is also seen in Microsoft Office.

1.1.2 Infrastructure-as-a-service (IaaS)

Infrastructure-as-a-service is an on-demand service delivery method that provides operating systems, servers, and storage via IP-based connections. Through an on-demand, outsourced service, clients may get these resources without having to purchase servers or software. Well-known examples of IaaS systems include IBM Cloud and Microsoft Azure.

1.1.3 Platform-as-a-service (PaaS)

Out of the three levels of cloud computing, it is said to be the most complicated. While PaaS and SaaS are similar, the main distinction is that PaaS is a platform for developing software that is supplied over the Internet, as opposed to software that is delivered online. This approach incorporates Heroku and Salesforce.com.

1.2. Cloud Disaster Recovery

Using cloud disaster recovery can help an organization become more resilient to calamities. Whether key workloads are kept in the cloud, on-premises, or in hybrid or multi-cloud settings, it helps safeguard them. Strong cloud disaster recovery (DR) solutions reduce downtime and related costs while assisting organizations in responding more effectively to cyber attacks and other calamities. Business continuity is made possible by fog computing, which guarantees the security and availability of crucial data, systems, and apps.

1.2.1 Cloud Computing

Fog computing is a decentralized computing method or architecture where computer resources are situated in between the cloud or any other data centre and the data source.

1.2.2 Fog computing vs. edge computing

Standards are developed for every new technical idea and serve as guidelines or instructions for users while using these concepts. When it comes to edge and fog computing, edge computing is the process of moving computers closer to data sources, while fog computing is a standard that outlines how it should be used in various contexts. Cisco invented the term "fog computing," which allows for consistency when using edge computing in a variety of industrial applications or tasks. Because they work together to lower processing latency by putting computers closer to data sources, they may be compared to two sides of a coin.



Figure 1.1. Fog Computing

2. System Architecture – VirtualBox Data Backup

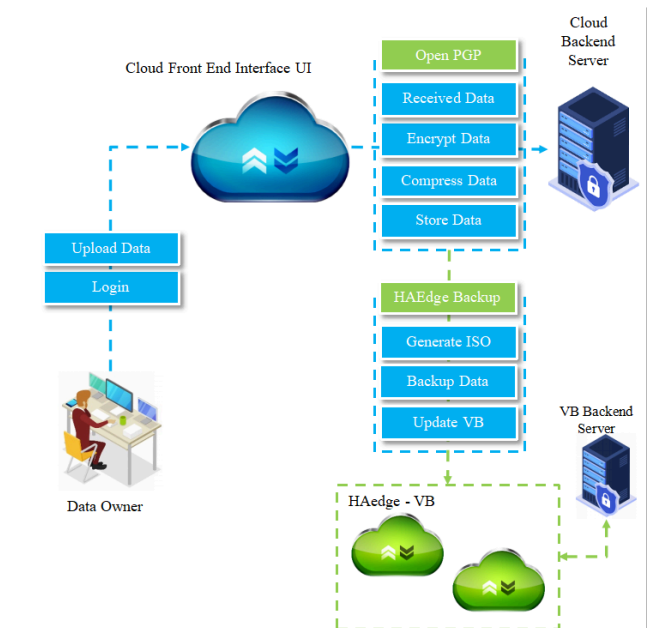


Figure 2.1. VirtualBox Data Backup

2.1 System Overview

The project involves developing a cloud-based application that integrates local and cloud-based elements. This programme is dependent on distant servers that are reachable via web browsers with a continuous internet connection. servers for cloud storage,

Through the help of independent cloud service providers, virtual storage facilities are made available online. These servers, which are housed in distant data centers, guarantee constant upkeep and functioning while protecting files from any harm. Application programming interfaces (APIs) make it easy for authorized users to access data stored on cloud services. These APIs make it possible for the cloud application and outside data sources or storage providers to communicate seamlessly.

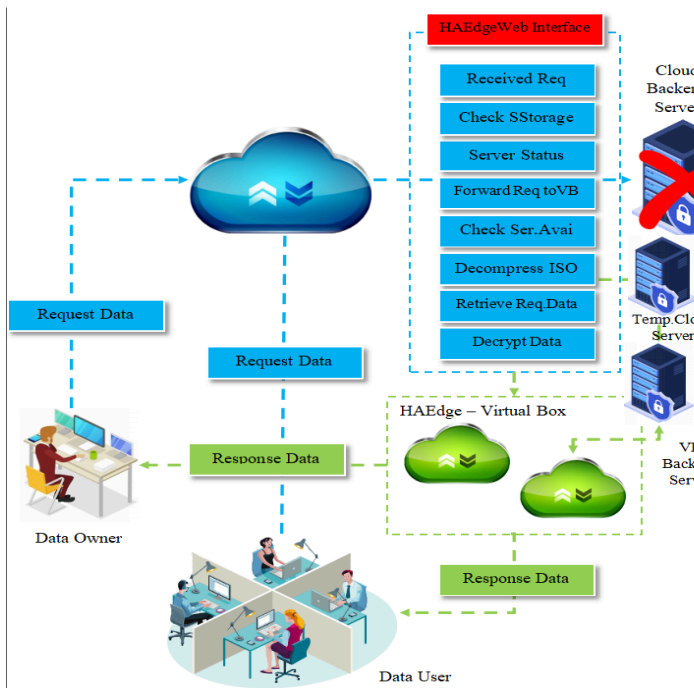


Figure 2.2. HAEdge Data Recovery

By simplifying the development process, the usage of APIs increases the application's predictability and efficiency. Three main categories of cloud storage are described in the framework: Block, file, and object storage are available. Systems such as Amazon Simple Storage Service (S3) exhibit scalability and metadata features that make object storage ideal for modern applications and analytics computations. File storage, offered by services like Amazon Elastic File System (EFS), may support applications that require a file system and shared files.

Block storage is designed for high-performance workloads and is frequently required for corporate applications like databases or ERP systems. It is exemplified by services like Amazon Elastic Block Store (EBS). The experiment also highlights how important home directories are to cloud processes.

In addition to supporting the option of keeping database backups in the cloud for short-term protection during upgrades or for development and testing needs, the project tackles issues associated with growing capacity requirements.

Despite its many advantages, cloud data migration recognises the worries of conventional IT functional owners, including those about networking, storage,

backup, security, and compliance. VirtualBox functions as an archive storage option within the project's structure.

The project's framework includes VirtualBox as an archive storage option. VirtualBox is a cost-effective solution that may be utilized for data storage for a minimum of 365 days. It is optimally installed on a dedicated server or any available spare local workstation, hence optimizing project productivity. Compression and data security are essential parts of the project, and data in transit is protected using the OpenPGP Protocol. PGP keys, both public and private, are used in the encryption and decryption procedures to guarantee the security of data transmission. Compression reduces the amount of plaintext and improves overall security, further optimizing the efficiency of data transmission. The cloud end user market is made up of people who depend on cloud servers for computing and data management, either alone or in groups.

VirtualBox is a feature of the project framework that allows for archive storing. VirtualBox is an affordable option that may be used to store data for a minimum of one year. It is set up to run best on a dedicated server or any spare local workstation that is available, which maximizes project productivity. The OpenPGP Protocol is used to secure data while it is in transit, making compression and data security crucial components of the project. To ensure the security of data transmission, PGP keys—both public and private—are utilized in the encryption and decryption processes. Compression maximizes data transmission efficiency by lowering the quantity of plaintext and enhancing security generally.

2.2. Modules Description

2.2.1. Cloud Service Provider Dashboard

We create a cloud application in this subject. A software programme that integrates local and cloud-based components is called a cloud application. In this notion, logic that is processed by remote servers is accessed through a web browser that is always connected to the internet. Cloud application servers are often situated in remote data centres run by a third-party cloud services infrastructure provider. Cloud-based apps may be used for file sharing, email, and storage, among other things. Third-party data sources and storage services can be accessed via an application programming interface (API). Cloud applications may stay smaller by sending data to apps or API-based back-end services for processing or analytics computations, and then

giving the findings back to the cloud application. Passive consistency is enforced by tested APIs, which can speed up development and yield reliable results. Data kept on cloud services is quickly accessible to authorized users.

2.2.2. Cloud Storage Server

Cloud storage servers are virtual locations for data storage that cloud service providers provide. They allow users to store and retrieve large amounts of data without requiring a physical device. An online storage server may be accessed over the internet. File Cloud offers cloud storage at an affordable price with no outages. The cloud service providers are in charge of maintaining these data centers, which allow the cloud storage servers to run continuously. Data centers store your information safely and ensure that you may view them online whenever you need to. Applications can directly access cloud storage using standard storage protocols or an API. Block storage, file storage, and object storage are the three categories of cloud data storage. Everyone has its own benefits and applications.

Object Storage:

Object storage's wide scalability and metadata characteristics are often leveraged by cloud-based applications. Existing data stores may be imported into object storage 2 systems, such as Amazon Simple Storage Service (S3), for archiving, backup, and analytics purposes. They are also ideal for creating modern applications from the ground up that require flexibility and scale.

File Storage:

A file system is necessary for certain programmes that need to access shared data. Network Attached Storage (NAS) servers are frequently used to support this kind of storage. Use cases such as music stores, user home directories, development environments, and huge content repositories are well suited for file storage systems such as Amazon Elastic File System (EFS).

Block Storage:

Other corporate applications, including databases or ERP systems, may require dedicated, low latency storage for each server. This is similar to direct-attached storage (DAS) or a Storage Area Network (SAN). Every virtual server comes with block-based cloud storage choices, such as Amazon Elastic Block Store (EBS), which offer the incredibly low latency required for high performance applications.

2.3. Home directories

For many cloud operations, it is helpful to utilize home directories to store files that are only accessible by designated users and groups. Many businesses are allowing their customers to access home directories in order to benefit from the cloud's scalability and cost advantages. Customers may quickly lift and transfer programmes that require this functionality to the cloud as cloud file storage solutions follow conventional rights models and necessary file system semantics.

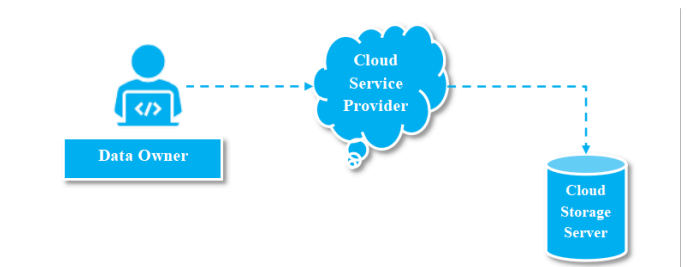


Figure 2.3. Home directories

2.4. HAEdge Application Layer

For Windows XP, Vista, and 7, HAEdge is a free and open-source computer imaging solution that integrates a few open-source technologies with a PHP-based online interface. As long as the data is under 40GB, HAEdge can also transfer an image from a PC with an 80GB partition to a device with a 40GB hard disc. Additionally, Fog comes with a Web Dashboard service that can be used to modify the PC's hostname, restart it in the event that a task is generated for it, and automatically import hosts into Cloud Storage Server and VirtualBox databases.

2.4.1 Backup and Recovery

Maintaining backups and recoveries is essential to making sure data is safe and readable, but it may be difficult to keep up with growing capacity needs. Existing software, processes, and semantics for data backup might result in isolated recovery scenarios with limited locational flexibility for recovery. Many companies wish to benefit from the flexibility of keeping database backups on the cloud, whether for development and testing purposes or as a temporary safeguard during upgrades. Cloud file storage systems can be a great platform to produce portable database backups using native application tools or business backup programmes, as they offer a standard file system that can be simply mounted from database servers.

2.4.2 Cloud Data Migration

Business owners may find cloud storage's affordability, dependability, and accessibility to be highly alluring. However, traditional IT functional owners, such as

networking, storage, backup, security, and compliance administrators, may be wary of the realities of moving substantial volumes of data to the cloud.

Archive Storage: Ideal for data that needs to be kept for a minimum of 365 days, such as regulatory archives. VirtualBox is best implemented on a dedicated server, any spare machine that company have locally. VirtualBox provides a feature-rich environment for creating and managing virtual machines (VMs), supporting key features like snapshotting for easy rollback and multiplatform compatibility. Creating VMs in VirtualBox is intuitive, thanks to its wizard-driven approach and the inclusion of Guest Additions – additional drivers and utilities enhancing guest operating system performance and integration. The software leverages hardware virtualization capabilities, such as Intel VT-x and AMD-V, to improve overall performance. Networking capabilities are flexible, offering modes like NAT, Bridged, and Host-Only, alongside port forwarding options for tailored network configurations. VirtualBox excels in storage management, supporting various disk image formats and dynamic allocation to optimize storage efficiency. Users can monitor resource usage in real-time, with the ability to set execution caps on CPU usage. The software is extensible, supporting extension packs for additional functionalities and APIs for automation and integration. The VirtualBox community is active, providing forums and resources for troubleshooting, complemented by comprehensive documentation. Widely used in software development, testing, education, and training, VirtualBox's versatility, user-friendly interface, and extensive feature set make it a popular choice for virtualization needs.

Top of Form

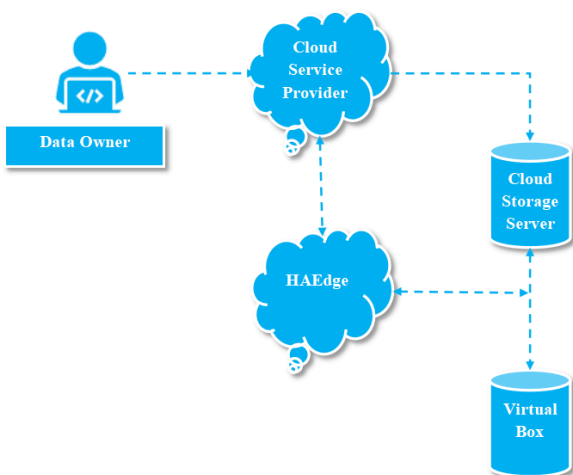


Figure 2.4. VirtualBox

2.5. Data Security and Compression

This module uses the OpenPGP Protocol to compress and encrypt data. OpenPGP is an open standard that is extensively used to protect cloud data, email services, and many other kinds of applications. Its goal is to safeguard data in transit by enabling end-to-end encrypted communication. Currently, OpenPGP supports contemporary cryptography and is regularly examined by well-known security specialists.

2.5.1. Encryption and Decryption

A key element of OpenPGP is data management. Your Data is separated between two public and private PGP keys, to put it simply. The private key is kept in the keychain of your device and has to be kept secret. Your correspondents can send you encrypted files that require your private key to decipher, provided you share the public key with them. Unfortunately, with this configuration, an attacker with access to your private PGP keys can surreptitiously assume your identity, accessing your encrypted communications and producing legitimate signatures. Applications still require access to these private keys in order to perform some PGP functions, even though you should keep an eye on their use and safeguard them.

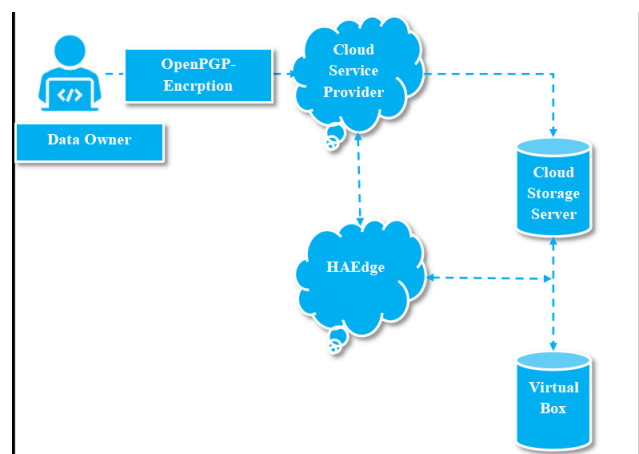


Figure 2.5. Encryption and Decryption

2.5.2. Zip Engine

Because plaintext is huge, it consumes needless disc space and modern transmission time. The user's plaintext will be compressed using a PGP programme to improve process efficiency overall. Additionally, this lessens plaintext patterns, which makes it more difficult for hackers to decode. Faster than ever ZIP engine Our sophisticated multi-core ZIP/ZIPX engine has been fine-tuned for optimal speed; it can now compress data at a similar strength while operating up to 30-50% quicker than the multi-core engine in WinZip (and significantly faster than the ZIP engines in SecureZIP and WinRAR).

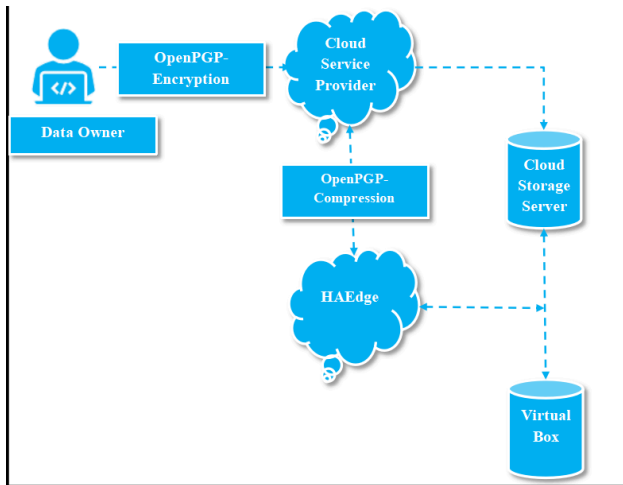


Figure 2.1.Zip Engine

2.6 Cloud End User

A person or group that uses cloud computing to offshore massive data files is referred to as a cloud user. These users totally rely on cloud server storage provided by cloud server providers (CSP) for compute and data upkeep. It can sustain and meet the demands of cloud users thanks to its extensive processing infrastructure and storage capacity.

2.6.1. Data Owner

data holder, who is the owner of the data and stores it at CSP in numerous blocks. It's possible that several qualified data holders in the CSP have access to the same encrypted data blocks. Specifically, with respect to those same blocks, the data owner is identified as the data holder who initially uploads the data blocks to the CSP. Before outsourcing, the data owner must create a dictionary W made up of unique terms taken from document collection D . This will ensure data privacy is protected while data availability is preserved. An unencrypted index tree may then be built using the dictionary and document collection. Lastly, the data owner outsources the encrypted versions of the index tree and document collection to the cloud server.

2.6.2 Data User

The owner of the data might grant the data user permission to see a certain document. The data user can create a trapdoor T with t query keywords in light of search control methods, and once the trapdoor is uploaded to the cloud server, k encrypted documents will be retrieved. Lastly, the data user can decrypt documents using the shared secret key. Authorities provide the data user with a set of characteristics and matching decryption keys. She/he is able to confirm if they are available and lawful. Any encrypted data can be freely obtained from CSP by the data user, who can only decode the ciphertext if certain requirements are met by her/his characteristics.

2.6.3 Performance Evaluation

Performance evaluation of the system to ensure its efficiency and responsiveness. The evaluation focuses on key metrics that directly impact user experience and system reliability. Here's a detailed description of the performance evaluation criteria:

3. PROPOSED WORK :

The suggested system presents HAEdge, a brand-new fog computing-based data backup solution. This system uses the benefits of Fog-Cloud storage to guarantee the security and dependability of users' data while simultaneously resolving multi-Cloud issues through the use of the Fog Computing paradigm. Without having to worry about the complex procedures required to safeguard and secure the data on multi-Cloud storage, system users may simply and securely backup, restore, and edit their data. suggested VirtualBox as a user-friendly, incredibly safe, and dependable backup solution that makes use of cutting-edge cloud and encryption technologies.

3.1 VirtualBox:

It's an individual Fog node. Like other personal devices like laptops and cellphones, it is owned and operated by the end user. One may think of VirtualBox as a Private Fog gadget. This bears some resemblance to the Private Cloud model, in which enterprises set up and maintain their own cloud infrastructure. The location of VirtualBox is in the network architecture. The Fog Computing paradigm may be used by several apps on the VirtualBox, a personal device. This feature enables VirtualBox to offer an improved backup experience as well as a unique solution that works better than the competition.

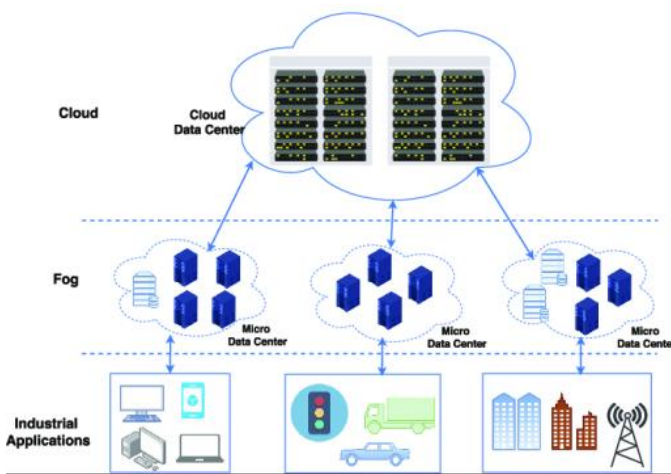


Figure 3.1. FogDrive

3.2 HAEdge:

HAEdge offers a distinct architecture for data backup systems. The advantages of multi-cloud computing and fog networking come together to create this uniqueness. Fog computing is utilized to give higher throughput and reduced latency for the backup process, while multi-cloud technologies are employed to provide the most dependable and secure storage environment. The VirtualBox fog device makes this special architecture possible. The client receives an easy-to-use dashboard to manage their data in VirtualBox and the cloud, all handled by the HAEdge provider.

3.2.1 OpenPGP:

Pretty Good Privacy (PGP) (OpenPGP) Plan-based With password protection, encrypted backups provide you personal control over your data, limiting access to only those who are authorized. Files can be moved to a remote server or connected to the network after being encrypted with PGP. It is possible to automate this entire procedure. A project workflow may be set up to, for instance, automatically extract data from a database, compile it into a CSV file, encrypt it, and transfer it to the SFTP server of your trade partner—all without requiring any human intervention from you or your team.

3.2.2 Open PGP Key Management:

PGP public and private keys are supplied along with an extensive Key Manager. It is possible to import, inspect, modify, and create keys with this key manager. You may use these keys in Cloud - Fog Data to automate PGP encryption and decryption for your company. Public keys may be exported from this Key Manager and shared with your trading partners.

3.2.3 Encryption:

A key pair is needed to carry out encryption. The receiver may receive the public key or it may be published. Only the receiver, who will decode the key, is aware of the private key. The file or communication is encrypted using the public key, and it is decrypted using the private key.

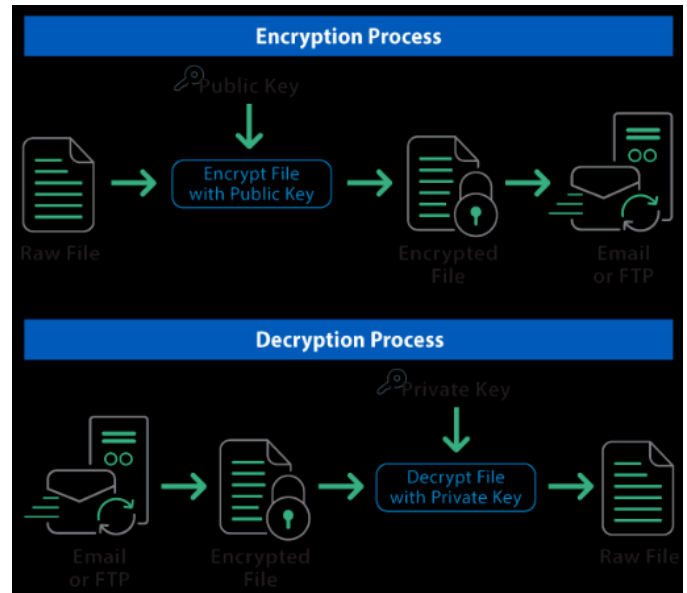


Figure 3.2. OpenPGP

4 RESULTS AND DISCUSSION:

The system's performance during a cloud storage server disaster, by the HAEdge App Layer and VirtualBox, reflects a strategic approach to ensure swift data recovery and service continuity. Early detection of anomalies triggers the activation of the HAEdge App Layer, showcasing the system's readiness for disaster recovery.

The reliance on VirtualBox for archive retrieval underscores the importance of maintaining essential system images and configurations, ensuring critical data accessibility during the recovery phase. The initiation of data restoration from backups, facilitated by HAEdge, emphasizes the system's commitment to preserving data integrity and leveraging redundancy to safeguard against potential data loss. The parallel activation of redundant systems by HAEdge ensures continuous data availability, minimizing downtime and ensuring uninterrupted service operation.

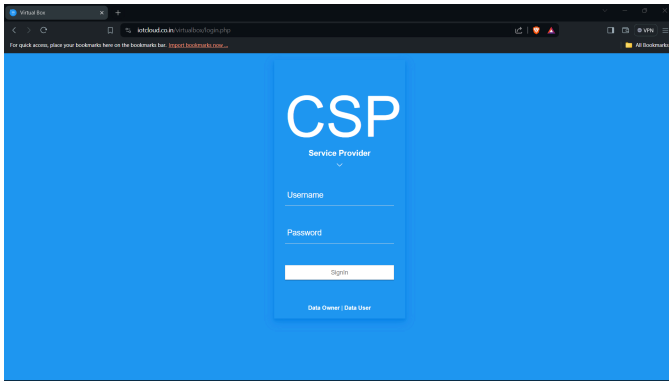


Figure 4.1. CSP

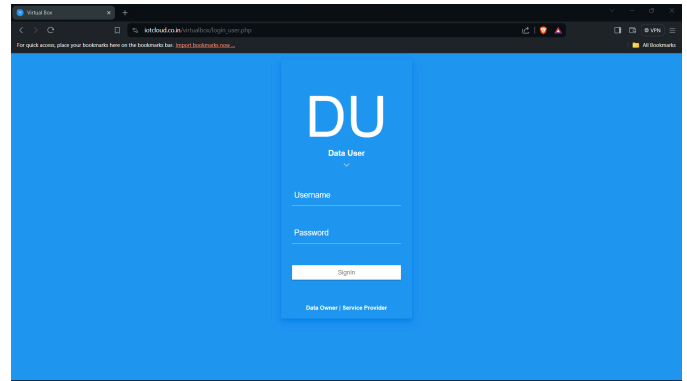


Figure 4.4. Data User Login Page

CSP- cloud service provider login page is used to approve and reject the data owner and they can upload the files into the cloud.

Data users can login to the website and they can upload the data.

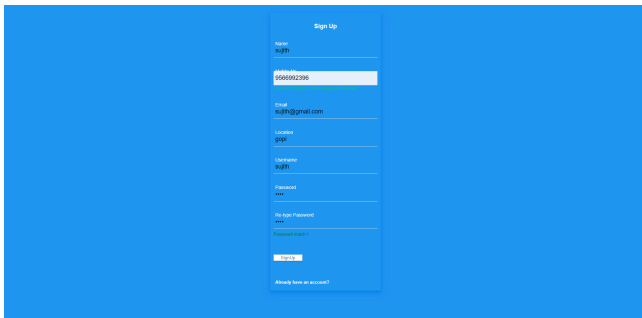


Figure 4.2. Data User sign-up page

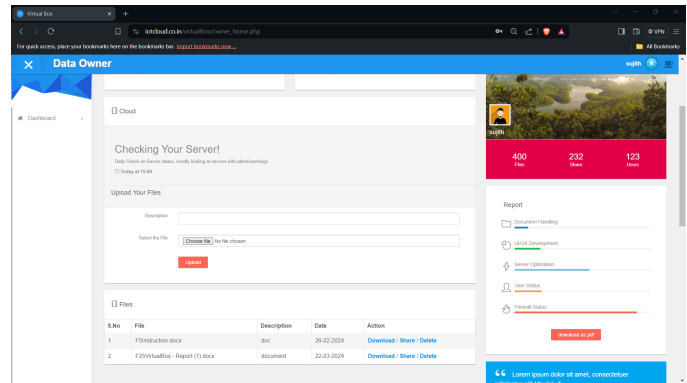


Figure 4.5. File upload

Data users can create the account after the CSP is approval

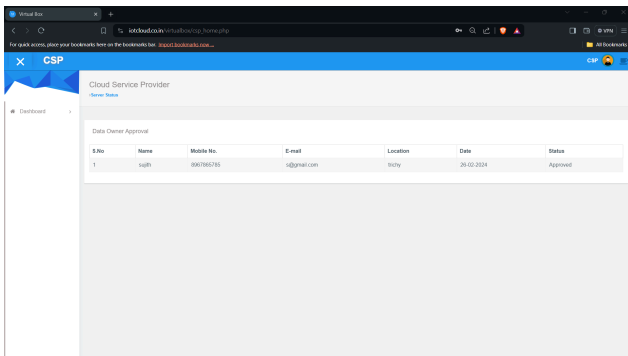


Figure 4.3. Data User approval

Data user approval page. it was made by cloud service provider

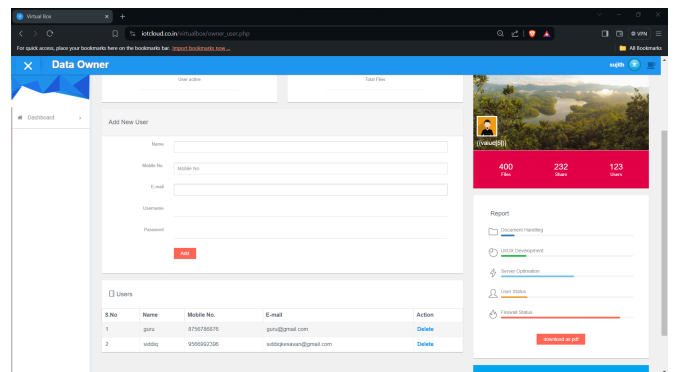


Figure 4.6. Data user creating

In this page we can create the data user login detail and share the files.

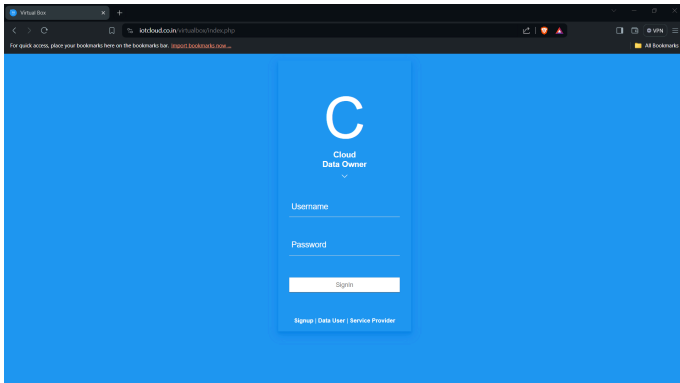


Figure 4.7. Data user login

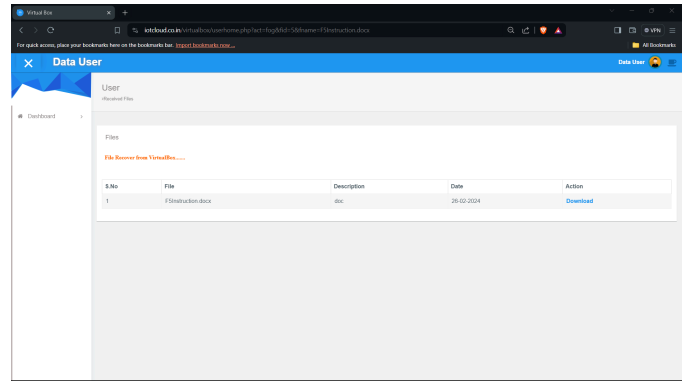


Figure 4.10 File recover from VirtualBox

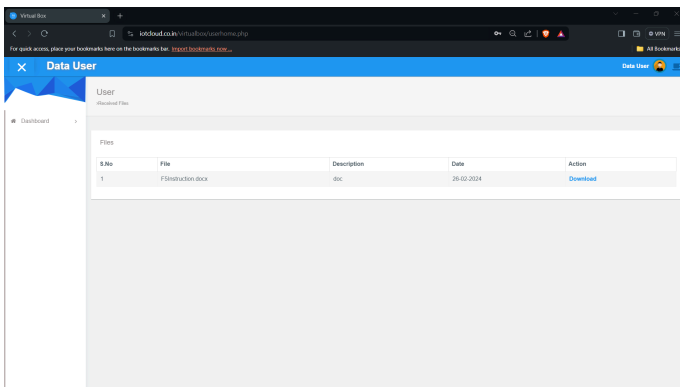


Figure 4.8. files downloading

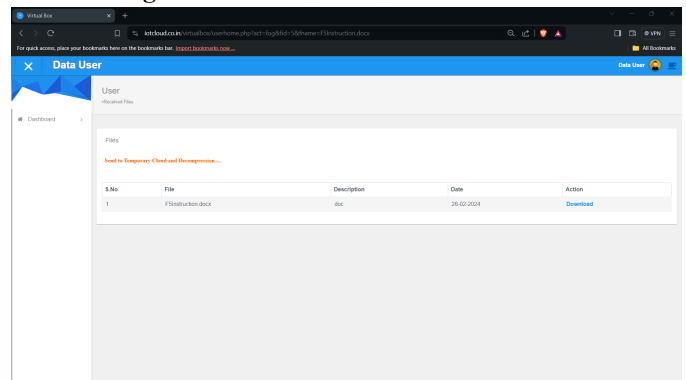


Figure 4.11. Temporary cloud and decompression

After they login to the website they can download the file. if the server is down they can't download the file.

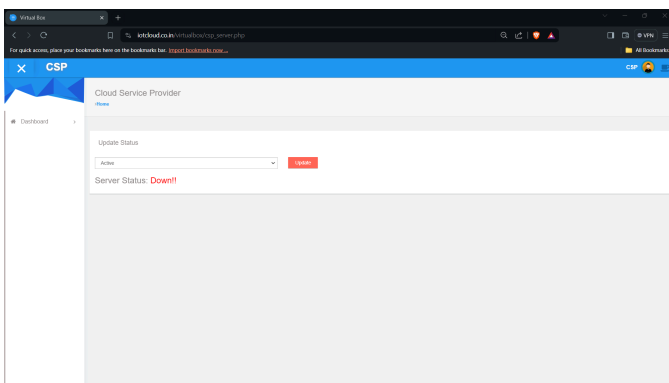


Figure 4.9. Server down

manually down the server to check if the HAEdge backup system is enabled

Granting temporary data access to users through alternative channels, facilitated by HAEdge and VirtualBox collaboration, highlights a user-centric approach, maintaining productivity during the recovery process. Real-time communication with users transparently provides updates on the recovery process, managing user expectations and building trust in the system's capabilities. Continuous monitoring, evaluation, and swift issue resolution demonstrate the system's adaptability, ensuring that any deviations from the recovery plan are promptly identified and addressed. Data verification and integrity checks post-restoration emphasize the system's commitment to maintaining the accuracy and reliability of restored data. The gradual resumption of normal operations ensures a controlled transition, minimizing potential disruptions and ensuring a smooth return to full operational status. Post-disaster analysis and documentation of lessons learned contribute to continuous improvement, informing future enhancements to the system's resilience against unforeseen challenges. Clear communication signals the successful resolution to users, managing expectations, and confirming the restoration of normal operations. The iterative update of the disaster recovery plan based on gained insights reflects a proactive and adaptive approach, ensuring the system remains well-prepared for potential future disasters. In essence, the system's performance during a disaster exemplifies a well-coordinated, user-centric, and continuously improving strategy for data recovery and service restoration.

5.CONCLUSION:

In conclusion, a strong and proactive strategy for data protection and service continuity is presented by the system's integration of the HAEdge App Layer and VirtualBox in the event of a cloud storage server disaster. The methodical sequence of events, from early identification to post-recovery analysis, highlights the system's flexibility and resilience in the face of unforeseen difficulties. The early detection method and the HAEdge App Layer's quick activation show how disaster recovery-ready the system is. Leveraging VirtualBox for archive retrieval ensures the availability of critical data, emphasizing the importance of maintaining essential system images and configurations. The initiation of data restoration from backups, guided by HAEdge, underscores the system's commitment to data integrity and redundancy. The parallel activation of redundant systems and the facilitation of temporary data access prioritize user-centricity, maintaining productivity during the recovery phase.

Real-time communication with users ensures transparency and manages expectations, fostering trust in the system's capabilities. This solution solves the issues with multi-Cloud by using the Fog Computing paradigm while also taking use of the benefits of temporary cloud storage to guarantee users' data security and dependability. Without having to worry about the complex procedures needed to safeguard and secure the data on temporary cloud storage, system users may simply and securely backup, recover, and edit their data. Comprehensive numerical outcomes show how well the suggested plan works to increase the survival of data and services in cloud data centers. This project can help data center operators strike a balance between data backup and service migration given a set of predetermined resources and early warning time limits.

References

1. Shivang Modi, Yash Dakwala and Vishwa Panchal, "Cloud Backup & Recovery Techniques of Cloud Computing and a Comparison between AWS and Azure Cloud", International Research Journal of Engineering and Technology (IRJET), vol. 07, no. 07, July 2020.
2. Abedallah Abualkishik, Ali Alwan and Yonis Gulzar, "Disaster Recovery in Cloud Computing Systems: An Overview", International Journal of Advanced Computer Science and Applications, vol. 11, pp. 702, 2020.
3. A.A.Tamimi, R. Dawood and L. Sadaqa, "Disaster Recovery Techniques in Cloud Computing", 2019 IEEE

Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), pp. 845-850, 2019.

4. S. Anuprabha and M. Nivaashini, "Protection of Cloud Services from Disaster Using Recovery Mechanism with Openstack", 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 1382-1387, 2018.

5. J. Mendonça, R. Lima, E. Queiroz, E. Andrade and D. S. Kim, "Evaluation of a Backup-as a-Service Environment for Disaster Recovery", 2019 IEEE Symposium on Computers and Communications (ISCC), pp. 1-6, 2019.

6. S. Hamadah and D. Aqel, "A proposed virtual private cloud-based disaster recovery strategy", 2019 IEEE Jordan Int. Jt. Conf. Electr. Eng. Inf. Technol. JEEIT 2019 - Proc., pp. 469-73, 2019.

7. Bhalerao and A. Pawar, "Utilizing Cloud Storage for Big Data Backups", pp. 933-938, 2018.

8. M. S. Fernando, "IT disaster recovery system to ensure the business continuity of an organization", 2017 Natl. Inf. Technol. Conf. NITC 2017, vol. 2017-Sept, pp. 46-8, 2018

9. Mohammad Alshammari and Ali Alwan, "A Conceptual Framework for Disaster Recovery and Business Continuity of Database Services in Multi- Cloud", 2017.

10. Raigonda rani Megha and Tahseen Fatima, "A Cloud Based Automatic Recovery and Backup System with Video compression", International journal of engineering and computer science., vol. 5, pp. 17819-17822, 2016