IJTBI
International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

CHARON INTELLIGENT MULTI-CLOUD SECURITY FOR SAFE DATA STORAGE AND COLLABORATION

[#1]**G. LAKSHMI,** ***Associate Professor,***
[#2]**BANGARU ANUSHA,** ***B.Tech Student,***
[#3]**PEDDAPELLI VARSHINI,** ***B.Tech Student,***
[#4]**KARAM ASHRITHA,** ***B.Tech Student,***
[#5]**ANDE SURABHISUBHIKSHITH,** ***B.Tech Student,***
[#6]**ERRAM SRICHANDANA,** ***B.Tech Student,***
Department of AIML,
TRINITY COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY, TG.

ABSTRACT: CHARON is a cloud-based storage system that is presented in this paper. Its purpose is to adhere to regulatory regulations regulating the security of private information while safely transferring and storing massive volumes of data across a variety of cloud providers and storage repositories. Scalability, decentralization (which eliminates the requirement for centralized servers), and the efficient handling of considerable data among a number of geographically dispersed storage providers are the defining characteristics of CHARON. It is possible for multiple clients to simultaneously attempt to access the same resources, which can result in write-write conflicts. In order to solve this problem, a one-of-a-kind lease method that makes use of Byzantine resilience was developed. The evaluation of CHARON is carried out using micro and application-oriented benchmarks that accurately imitate the procedures that are used in the field of bioinformatics. The findings indicate that our innovative design outperforms other cloud-based systems by a factor of 2.5 across all criteria, which indicates that our system is more functional and performs better.

Keywords: Big-data storage, Cloud storage, Byzantine fault tolerance.

1. INTRODUCTION

"Cloud storage" refers to the storage of information on remote servers and databases that are accessible over a network, typically the Internet. The term "cloud computing" originated from system diagrams that depicted a cloud icon for the backend configuration. Remote providers are trusted by users to manage their data, applications, and processing when utilizing cloud storage. "Cloud storage" is a distributed system of computers and software that is used to store and retrieve data. These components facilitate the operation of contemporary applications and the access of modern networks by servers.

Figure 1: Cloud Backup Services

The optimum environment for high-throughput computing is the cloud. Applications aimed at consumers, such as financial portfolios, are its intended usage. Such apps may provide users with personalized content, retain their data, and

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

create entertaining online games. On a daily basis, both the military and research labs use sophisticated computer systems.

Cloud storage networks are comprised of groups of interconnected computers that are linked to the internet. These groups of computers, which are often inexpensive personal computers, pool their processing power to handle large amounts of data. Connectivity between several large networks is made possible by the IT architecture that was previously described. Cloud computing is made even better and faster by using virtualization techniques.

Versions of Features and Services:

The phrase "cloud computing" and its most important components are defined below using NIST-created terminology.

On-demand self-service:

A client can easily and rapidly make available, on demand, one-way computational resources like processing power and network storage without contacting each service provider individually.

Wide network access:

Standardized protocols allow users to connect to a network from a greater variety of client systems, including PDAs, laptops, and mobile phones.

Capital pooling:

The provider employs a multi-tenant method, which involves pooling its computer resources and making them available to multiple clients simultaneously. This method's physical and virtual components are independent. Appropriate resource management includes allocating and reallocating resources in response to changing consumer needs. In most cases, users aren't privy to or given much say over where exactly the available facilities are located. They might just consider the location at the national, state, or data center levels. Among the many things you can acquire are storage, processing power, random access memory, CPU speed, data transfer rate, and virtual machines.

Rapid elasticity:

Scalability refers to the simplicity and rapidity with which capacity can be raised or lowered to meet changing demand. Many consumers mistakenly believe they have unlimited access to these provisioning resources.

Measured service:

Storage, processing, bandwidth, and active user accounts are all measurable metrics. Various levels of abstraction can be measured with the help of cloud providers. There needs to be a system in place to monitor and record service consumption in order to hold service providers and consumers accountable.

2. RELATED WORK

The CHARON cloud storage system mimics the POSIX protocol. The objective is to simplify the process of storing and sharing large datasets without requiring expertise in complex software or operations. The need for a more efficient method of managing genetic data prompted the development of this device for use in the bioinformatics and biodiversity industries.

The CHARON system encrypts documents, uses erasure codes and compression algorithms to make them smaller and simpler to handle, polishes

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

them, and references information from previous uploads in addition to daily processing of massive amounts of data. The unique form and an assortment of characteristics of CHARON distinguish it from the throng. The reason is that it is a convenient tool that consolidates numerous concepts that were previously distinct. CHARON surpasses competing multi-cloud solutions by a factor of two to four throughout the entire process.

The speed at which it processes files means that its correctness is comparable to that of the popular NFS (Network File System). Making CHARON, a cloud-connected gadget that facilitates the transfer of massive volumes of data, is the primary objective of the project. The device is safeguarded against Byzantine attacks by using a data-centric leasing technique. Our algorithm integrates the finest features of multiple cloud providers rather than depending on the popularity of a single one. A thorough evaluation of CHARON's functionality requires testing on several storage systems, including local, networked, and cloud-based options. The input and output performance of bioinformatics software is tested using micro benchmarks and a domain-specific benchmark.

3. SYSTEM DESIGN

The system's CHARON distributed file system facilitates user access to the cloud and data sharing amongst users. Both the POSIX standard and the CHARON interface are very similar. Due to the fact that the majority of current life sciences solutions rely on file-based inputs and that the intended users are generally not professionals, the decision was made to utilize a POSIX interface rather than data objects.

The technology's stated goal is to facilitate the secure sharing of files, the management of several storage facilities, and the storing of massive amounts of data. In order to make deployment as easy and fast as possible, we aim to eliminate user-deployed servers and make minimal changes to existing cloud services.

The situation is exacerbated by this. With the completion of two crucial design decisions, everything in CHARON fell into place. The files are incrementally transferred to their final destination after being written to the client's local storage. Prefetching and parallel downloading are common methods used to speed up reading. Consequently, CHARON's value increases. This choice is affected by the file's size and its intended recipient. In particular, (1) customers are unlikely to be knowledgeable about dispute resolution, (2) manually resolving conflicts in big files can be tedious and time-consuming, and (3) keeping several copies of this data can be costly. Particularly for heavily utilized libraries, such as Google Genomics, these quality control procedures are crucial. The sample repository is where these libraries keep track of the results of data processing activities. This allows users to get new insights from the data and share them with the community.

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

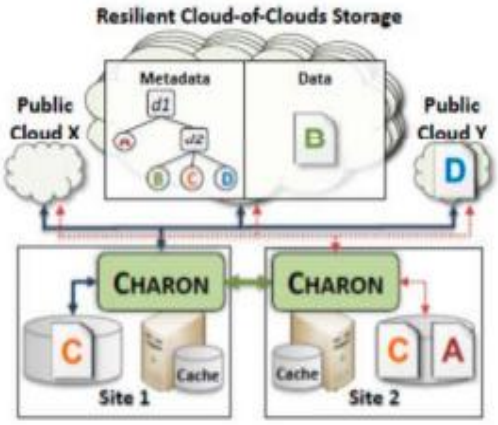

Fig2. System Architecture

SHA ALGORITHM:

Data privacy and security are the primary objectives of developing cryptographic features. Included in this broader category are Secure Hash Algorithms (SHA). Bitwise operations, modular additions, and compression features are all part of a hash function, which is used to reconstruct the statistics. After that, the hash function will generate a string of a specific length that closely resembles the initial value of zero. You can't utilize character hash values to retrieve the original statistical data because the methods discussed here only work in one direction.

SHA-1, SHA-2, and SHA-3 are among the most renowned cryptographic algorithms that were created to thwart the development of increasingly sophisticated hacking attempts. There is a general consensus that SHA-0 is no longer suitable for authoritative consensus due to its numerous documented and recognized deficiencies. FUSE-J, a Java interface that was created for the FUSE library, is employed by the CHARON utility to generate user interface documentation. The client will be responsible for all system administration duties, in addition to managing and storing data in the cloud. The procedure is also available as free software that can be downloaded.

Metadata Organization:

Data attributes, such as those of a file or directory, are examples of metadata. A single author and multiple readers are made possible by CHARON's storage system, which employs a cloud-of-clouds architecture with registers, regardless of the physical location of the data pieces. Both retrieval and use are made easier by this. Both the concurrency and presentation of DepSky improved following the redesign and optimization of the SWMR check.

Managing namespaces:

The majority of metadata is stored in namespace objects. They display the structure of the subdirectory tree, which includes both directories and documents. CHARON uses SNS and PNS, which stand for public and private namespaces, respectively. A Personalized Notification System (PNS) stores the information on all of a buyer's non-collectible possessions. The quantity of social media sites that a buyer can access is directly proportional to the number of folders that it can access.

One social networking service (SNS) is linked to each purchaser's PNS for each shared container. A systematic approach to managing shared files that promotes effective user cooperation, accessibility, and organization is highly recommended (Section 4.1.2). The first step in accomplishing this is to arrange your files in a sensible manner. Improving your ability to think laterally is also important. The PNS data can be swiftly retrieved following the document device's setup using the cloud-of-clouds architecture.

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

Conversely, SNSs are keen on regularly receiving updated metadata from community directories. Client Y might pull off the most incredible stunts using the rented social media platform while Client X is busy writing.

Data Management:

CHARON uses this technique as one of its primary methods for organizing lengthy content. Without reliable service providers, it is next to impossible to receive helpful information. The data is encrypted and the keys are securely protected by secret sharing during the single-writer multiple-reader (SWMR) signup procedure. An abstraction mapping involving the document device and the cloud storage space must be established in order to maximize the utilization of this version.

The frequently used documents are stored in CHARON's local cache. To facilitate the retrieval of previously stored data, it maintains a tiny, fixed cache in main memory. Priority setting in caches is based on the LRU algorithm. When working with large documents in cloud-enabled document infrastructures, it becomes extremely challenging to use data and analytics. To start, the extremely high latency associated with uploading and receiving large text files from cloud storage makes it difficult to access or create such files. Never forget that cloud-based document management solutions could struggle to handle a large number of large documents in their memory cache. An astute approach to the issue at hand is CHARON's partitioning technique. It partitions large documents into 16 MB chunks of a predetermined size. This method's blocks, after undergoing solidity and erasure coding, can store many megabytes. The present version of CHARON allows the document owner to pay for storage and exercise ownership rights; it has received accolades for striking a good balance between latency and performance. All customers are expected to contribute fairly to the costs incurred when accessing shared directories, as stated in this policy. When it comes to modifying and assigning permissions for each item, CHARON customers have no idea how their cloud providers will handle it.

In addition, the admission control of the cloud-of-clouds would continue to be satisfied regardless of the actions of a single cloud provider. When you look at something from a specific angle, this effect takes place.

The precision of the query stop end result obtained utilizing the resources of the question-issuing node is shown on the vertical axis, and the number of statistical devices queried is shown on the horizontal axis. No matter how many data devices you utilize, the recommended pinnacle-ok query method should continue to function well. Figure 6 shows that site visits occurred in tandem with the display of the various query result formats. On the one hand, we have the audience demographics shown on the X-axis, and on the other, we have the many different statistical approaches that were used to examine the data. A maximum of three possibly harmful nodes can be located with significantly reduced query resource use, according to the hazardous node detection ratio. The question's publication date and

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

the total number of incorrect answers are displayed on the horizontal axis.

METHODOLOGY:-

The superior encryption standard for block ciphers is their ability to handle 128-bit blocks, regardless of whether the key length is 192 bits, 256 bits, or 128 bits. Additionally, keep the following in mind while you search for a new and improved encryption method that comes highly recommended and drastically improves security:

Security

The competition's emphasis on security didn't stop the selection of hard algorithms over other post-ciphers primarily on their invulnerability to hacking.

Cost:

Memory and computational efficiency tests were conducted on the prospective algorithms prior to their distribution with a worldwide, nonexclusive, royalty-free license.

4. IMPLEMENTATION

The algorithm's characteristics will be examined and comprehended in order to generate rules that are useful, adaptable, and compatible with both software and hardware. There are three distinct ciphers that make use of the Advanced Encryption Standard (AES), and they are AES-128, AES-192, and AES-256. Each of them has a different key length. All ciphers require cryptographic keys with lengths of 128 bits, 192 bits, or 256 bits in order to encrypt and decrypt data segments that are 128 bits in length. 128 bits is the minimum length required. Since its creation, Advanced Encryption Standard (AES) encryption has refrained from making use of the Rijndael cipher's adaptability so that it may accommodate a variety of block sizes and key lengths.

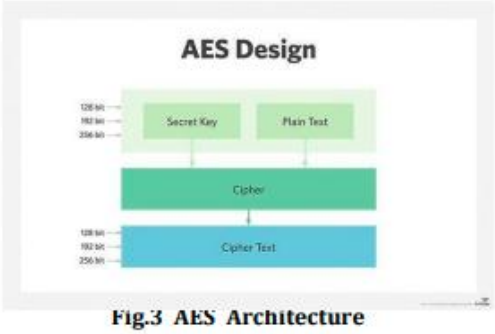

Fig.3 AES Architecture

5. CONCLUSION

Subsidized cloud services were the focus of this in-depth analysis of efficient submission procedures. Even users without additional compensation can benefit from the increased efficiency and speed of cloud-based submission techniques. Complicated data backup and restoration processes are a part of low-cost cloud replication. It provides a unified method for companies to secure their data. You can find a wealth of information there about the control services, power generation, disaster recovery plan, and cost-cutting measures offered by the organization. To facilitate the management and sharing of massive datasets, CHARON was developed as a cloud-based submission system.

A data-centric machine can be run on top of this system, which relies on the ease of syncing data and document metadata across several cloud platforms. Through the extension of a very precise Byzantine robust leasing protocol, we have successfully eliminated write-write conflicts from this system. A second server is unnecessary in this case. Our research's findings provide credence to the idea that

International Journal of Technovation and Business Insights
Volume 02, Issue 01, January-March 2026
ISSN: 3107-7900, P.CODE: 26/2-1/16
DOI: https://doi.org/10.5281/zenodo.19130388

this method may be useful in practical settings where sensitive information must be securely maintained and transmitted.

The CHARON file system, designed to store and share massive amounts of data, is supported by the cloud. By consistently distributing data, metadata, and files across various cloud platforms, this technique eliminates the need to have blind faith in any one cloud provider. Without the need for a dedicated server, our one-of-a-kind leasing system is Byzantine fault-tolerant and substantially reduces the likelihood of destructive write-write conflicts. Our findings provide credence to the practicality and possible benefits of the suggested method for protecting the transfer and storage of sensitive information.

REFERENCES

[1] Cloud Harmony, "Service Status," https://cloudharmony.com/status-of-storage-groupby-regions, 2019.

[2] Cloud Security Alliance, "Top Threats," https://cloudsecurityalliance.org/group/top-threats/, 2016.

[3] M. A. C. Dekker, "Critical Cloud Computing: A CIIP perspective on cloud computing services (v1.0)," European Network and Information Security Agency (ENISA), Tech. Rep., 2012.

[4] H. S. Gunawi et al., "Why does the cloud stop computing?: Lessons from hundreds of service outages," in Proc. of the SoCC, 2016.

[5] European Commission, "Data protection," https://ec.europa.eu/info/law/law-topic/dataprotection en, 2018.

[6] G. Gaskell and M. W. Bauer, Genomics and Society: Legal, Ethical and Social Dimensions. Routledge, 2013.

[7] A. Bessani et al., "BiobankCloud: a platform for the secure storage,sharing, and processing of large biomedical data sets," in DMAH, 2015.

[8] H. Gottweis et al., "Biobanks for Europe: A challenge for governance," European Commission, Directorate-General for Research and Innovation, Tech. Rep., 2012.

[9] P. E. Verissimo and A. Bessani, "Ebiobanking: What have you done to my cell samples?" IEEE Security Privacy, vol. 11, no. 6, pp. 62–65, 2013.

[10] P. R. Burton et al., "Size matters: just how big is big? Quantifying realistic sample, requirements for human genome epidemiology," Int J Epidemiol, vol. 38, no.1,pp.263–273