JCST Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776479

ADVANCED-DATA RECOVERY AND SECURITY USING CAUCHY CODING IN DISTRIBUTED CLOUD ENVIRONMENT

Mahesh Bhaskar Gunjal and Dr. Sonawane Vijay Ramnath
Department of Computer Science and Engineering,
Dr. A. P. J. Abdul Kalam University, Indore (M.P.) - 452016, India
Corresponding Author Email : maheshgunjal2010@gmail.com

Abstract- :In the modern era of cloud computing, the rate at which new data is produced and stored in the cloud is quite high. As a result, the cloud storage platform has emerged as one of the most important aspects of cloud computing. Cloud computing is typically used for storage systems these days due to its many desirable properties, including extensive network access, resource pooling, an on-demand approach, monitored service, and so on. Users' data may be lost if certain unfavourable situations arise, such as an electrical surge, a malfunctioning spindle motor, mechanical damage, or faults in the underlying programming. The data storage system ought to provide some kind of protection for the user's files. Replication is used to accomplish this goal. Although it is simple to store duplicated data to sustain a given amount of data loss, the storage performance is quite poor. Erasure codes are replacing replication in cloud file systems, which is being done primarily to improve data security while simultaneously cutting down on storage overhead. The erasure code approach is utilised to provide security for the data that is present across distributed systems. Erasure coding methods that are now in use, like reed-Solomon codes, seem to be able to significantly reduce the cost of data storage while maintaining the same level of resilience against disc failures. On the other hand, it has far higher costs associated with its repair, not to note an even longer access latency. From this perspective, the development of innovative coding strategies for cloud storage systems has attracted a substantial amount of attention both in academic circles and in the business world. Therefore, the objective of this research is to present a novel coding scheme that improves the effectiveness of the Reed-Solomon coding method that makes use of the advanced Cauchy matrix to accomplish fault tolerance.

Keywords: Distributed Storage System, Advanced Cauchy Reed Solomon, Erasure Code, Data Security, Fault Tolerance

I. INTRODUCTION

The amount of data that is created and saved in the modern world is growing at an exponential rate with time. Because an increasing number of people are connecting to the internet via various platforms, such as social media, blog posts, and online shopping, an enormous quantity of data is being pushed into the archives of these systems [2]. Because this data is so important to their operations, companies cannot risk losing even a single solitary piece of it. These companies conduct a great deal of research and assessment on this data, and it base their business decisions on the findings of these studies and analyses. This avalanche of information has reached unprecedented heights ever since the invention of cellphones. People can become more connected because of smartphones, and as a result, every photograph they take of a moment of ceremony in their lives, every visual or audio clip they enjoy, every humorous message that helps them laugh off their stress, and every important piece of data that they want to convey to a set of individuals are all uploaded to these portals [3,5].

At this time, the quantity of data that is created and saved regularly is on the scale of exa, and petabytes, and it is continuing to grow at an exponential rate. Disk failures are a regular concern for huge storage systems, and it is necessary to be tolerant of this fact to prevent the loss of data [11]. A "replication" of the disc was a common practice in the past for data protection. Because of the enormous magnitude of the data, it is no longer feasible to replicate the data in order to guarantee its availability. Because the default replication factor is three (a replica factor of three assures that it can sustain two failures at the same time), an additional storage overhead of 200% is required. Additionally, replication is responsible for the issue of consistency amongst replicas, which arises from the fact that active replication between the replicas is ineffective due to a variety of network-related issues [13,14,15].

These days, data grows at an exponential rate, which means there is a high probability of loss. As a result, there is a need for a system that supports the option to recover lost data within a reasonable amount of time [18]. Further on, erasure coding was also utilised as a way for achieving fault tolerance. However, it only produces a limited amount of tolerance, which is not sufficient when applied to data from actual applications. The likelihood of a disc failing in a distributed storage system is proportional to the size of the storage system. Therefore, the strategy of "Reed Solomon" can be utilised to cope with increased loss.

In this particular research study, the Advanced Cauchy Reed Solomon (ACRS) method serves as an enhancement to the 'Reed Solomon method. This accomplishes its goal of simplifying the process of encoding and preserving the erasure, as well as retrieving information after it has been lost, by utilising the XOR operation, which is capable of bypassing the Galois Filed Multiplication. This 'ACRS' code turns 'k' data blocks into 'm' coding blocks, and by employing the new strategy, it is possible to handle 'm' disc failure as well.

To accomplish this, the following steps were taken: first, a Cauchy matrix of data blocks was located, and then, after writing down every possible combination of the information matrix and scheduler utilised to encode, the best match pair was identified. However, as of right now, there is no such approach that will identify the best match because the priorities of the data rely on the user, and the redundancy of each piece of data that is required is unique. Therefore, the only choice that is left is to write down all of them, evaluate them, and pick the best one also conclusions can be drawn based on the results of experiments.

Even though erasure coding and RAID systems are linked to one another, erasure codes have some advantages over RAID systems. These advantages include the flexibility of fault tolerance to a range of circumstances, a smaller storage footprint for fault-tolerant in comparison to mirroring, and the capacity to scale in a distributed manner. Erasure coding is currently receiving a lot of attention from researchers because it could be a useful technology for cloud storage. A brief introduction to the idea of erasure coding is discussed as follows. The following is the format in which erasure coding is structured when a data set of size Q is provided. The data set Q is partitioned into p informational chunks with the same amount of detail in each. After that, the k data blocks are coded into n chunks of the same size as one another. The p data blocks and the n-p parity blocks are contained within the n blocks. The contents of these blocks are stored into n different storage nodes. The nodes of storage that hold data are called data nodes, and the nodes that contain parity bits are termed parity nodes. If any n-p nodes fail, the data that was lost can be recreated with the help of the p nodes that are still operational. New nodes, also known as newbie nodes, can recreate the data of failing nodes by obtaining the necessary blocks from p different live nodes.

With erasure-coded storage technologies, the reconstructing mechanism serves as the method for providing fault tolerance and, as a result, data durability. When the master server receives a file request from a client, it analyzes the pieces of the file by referencing the metadata, merges the contents of the file, and then sends the file to the client. The master will begin the process of data reconstruction to fulfil the client's request if all of the storage nodes that were holding the pieces failed at the same time. As opposed to what is referred to as "normal reads," these instances are referred to as "degraded reads." Erasure Coding gets its name from the fact that the coding makes it easier to recreate the data in a node that has been wiped (or failed). Erasure coding is a method that utilizes the use of the MDS codes' error-correcting ability. A Maximum Distance Separable (MDS) code is referred to as a linear code when it satisfies the bound ($d <= n - p+ 1$).

The proposed work aims to identify the optimum match because there are no combinations of the matrix as well as a scheduler that provides the greatest performance for all possible redundancy setups. This is because the replication factor of the data varies depending on the user prioritization of the data. In addition to this, it seeks to discover an improved coding scheme for the varying combinations of the matrix as well as the schedule. Since there is a significant disparity between the amount of XOR operations as well as the number of operations required to locate the schema for every combination, the system will complete its activities in the shortest possible amount of time and with the fewest possible number of operations.

The paper is fragmented into several sections. In section II, the theoretical background of the research is discussed. In section III, research carried out by various researchers is discussed more in detail. In Section IV, the proposed research methodology is explained in more depth. In Sections V and VI, experimental results and conclusion are discussed.

II. RELATED WORK

Cloud computing presents many challenges, one of the most significant being the protection of resources against access by unauthorized users, as stated by Soniet al.[1]. The role-based access control model, also known as the RBAC model, is a typical way of controlling data access. It is an access control that is extensively deployed in a variety of cloud systems. The level of complexity rises in proportion to the number of roles that are performed. The RBAC model has been replaced with the attribute-based access control model, which provides more scope for customization. In this instance, roles are linked directly to their associated permissions. The combination of RBAC with ABAC results in the development of a hybrid access control strategy that is both more extensible and much more dynamic. It begins with an overview of the various models of data access control and continues with a comparison of the properties shared by each. Comparisons are made between the various aspects of each of these security models.

Amitesh Singh et al. [2] came up with an original method for enhancing image colours that were cloud-based and protected users' privacy. It was proposed that colour correction operations be carried out in the encrypted domain over the cloud, which is in contrast to the colour correction schemes that are already in existence. In these schemes, the colours of the test pictures are analyzed in the plain domain alongside the visible image contents. As a direct outcome of this, greater achievements are accomplished with the guarantee of total confidentiality. The proposed method was put to the test using many different experiments, and it was proven to be highly efficient when compared to other methods that are considered to be state-of-the-art. As we move forward, the security resiliency of the suggested approach will be proved using a challenge-response strategy.

Nehad H. Hussein [3] proposed the use of a cloud-based system for the safe storage and distribution of medical pictures. The proposed system uses a variety of different cryptographic methods to provide increased safety for the medical data both while they are being transmitted and while they are stored in the cloud. In this context, the methods known as Elliptic Curve Cryptography, Advanced Encryption Standard and Secure Hash are utilized. A third-party auditor is utilized to validate the quality and validity of medical photographs prior to their being stored in the cloud. This helps to ensure that the clients' devices are not overloaded with too much processing. In addition, a digital signature is generated to guarantee the authenticity of the data source and to guarantee the robustness of the proposed method against any action that would disclose or alter the data. These two goals are simultaneously accomplished. The findings of the security analysis demonstrate that the technique ensures a high level of information security. The performance analyses have revealed that the system offers a high storage capacity while only introducing an acceptable amount of storage and computation overhead.

A secure cloud-based storage solution for an Internet of Things environment was proposed by Jeong, Junho et al. [4]. The service is built on a proven data possession paradigm and makes use of Bloom filters. Even if it generates false positives due to the employment of the Bloom filter, the outcomes of the experiments showed that the suggested technique is faster than the current methods and does not significantly differ from them in terms of the verification rate. As a result, the service that has been presented is able to analyze an impressive volume of data in an Internet of Things setting effectively.

Cui et al. [5] evaluate and improve the constraints of the conventional RBAC framework, and propose an upgraded RBAC model known as ET-RBAC as a solution to these drawbacks. On the base of the original RBAC paradigm, ET-RBAC adds the limitations of the environment component and time component to realize dynamic authorisation and resource management of roles. This is done by utilizing the modules' respective times and environments. When revoking resources and rights, the non-relevant revocation approach is used to prevent the privileges of the next-level role be influenced by the dynamic changes made to the roles above it. It created and applied the ET-RBAC model, which included designing roles, users, permissions, family groups, resource allocation mechanisms, and database management models. The ET-RBAC model is incorporated into the home automation management system, which it designed. Once the overall design has been finished, this model can be used for a variety of application platforms that have access control criteria to guarantee the safety of app resource availability.

Sukmana et al. [6] offer a uniform cloud access control approach that offers an encapsulation of the services provided by CSPs for centralized and autonomous administration of cloud resources and access control across multiple CSPs. The privilege isolation concept and the concept of least privilege are followed by the proposal to provide role-based access control for CSB stakeholders to access the cloud resources. This is accomplished by allocating necessary privileges as well as access control lists to cloud resources and CSB stakeholders, respectively. The unified model was implemented in a CSB system that was given the name CloudRAID for Business (CfB), and the assessment result indicates that it delivers system-and-cloud level security service for cfB in addition to centralized resource and access control management in different CSPs.

Meixia Yang and colleagues [7] developed a microservices-based OpenStack surveillance system, which they called OpenStack-reporter. The system's goal was to analyze OpenStack clouds and provide an easy-to-use tool for OpenStack managers. The OpenStack Reporter project has adopted a microservices design, and this architecture is comprised of three main components. Every component is accountable for a certain aspect of the project. The Kubernetes DNS-based service discovery and RBAC methods are responsible for implementing the connectivity between these many components. Kubernetes, which is mostly used for automated rollouts and rollbacks, is the system that is responsible for managing all three of these components. An OpenStack base and an OpenStack-reporter have each been constructed in the datacenter and the management centre, correspondingly, and the results are published. This was done so that the effectiveness of the proposed monitoring system could be verified.

According to the research of KritikaSoni et al. [8] In cloud computing, one of the most significant challenges is preventing unauthorized users from accessing the resources. The RBAC model is a typical way of controlling data access. It is an access control that is extensively deployed in a variety of cloud systems. The level of complexity rises in direct proportion to the number of roles that are played. The RBAC model has been replaced with the attribute-based access control model, which provides more leeway for customization. In this instance, roles are linked directly to their associated permissions. The combination of RBAC with ABAC results in the development of a hybrid access control strategy that is both more scalable and more dynamic. It presents the many different models of data access control and compares the properties of each of them. Comparisons are made between the various aspects of each of these security models.

Mahdi Ghafoorian et al. [9] suggest a new trust and reputation-oriented RBAC design. This model not only can appropriately withstand the vulnerabilities that trust-based RBAC frameworks present, but it is also extensible because its running time is acceptable. Third, the proposed model is analyzed by using the well-known dataset that contains the Advogato trust network. In the end, a comparison was made between the developed framework and published recently regarding a mean absolute error, the thecompletion time of implicit trust evaluation and given features. The findings that were achieved indicate the significance of the suggested model to be used in situations clouds.

The Intelligent RBAC (I-RBAC) model was developed by Rubina Ghazal et al. [10] as an innovative access control system that makes use of semantic business functions and expert systems. It comprises occupational privileges as positions for multiple disciplines. It makes use of the dataset of the initial occupational roles that were offered by the Standard Occupational Classification (SOC), which is located in the United States. The innovation of the study comes in the fact that it develops a core I-RBAC ontology by combining real-world semantic business roles as well as intelligent agent techniques. This is done to provide the requisite degree of access control in a fast-changing multi-domain context. To automate the generation of organisational ontology from unstructured format policies, the intelligent agents make use of WordNet as well as bidirectional long short-term memory deep neural networks. Further matching is done between this dynamically learnt organisational ontology and the basic I-RBAC ontology to derive uniform semantic business roles. A mathematical description of the suggested I-RBAC model, as well as an explanation of the general I-RBAC architecture and its operational architecture, are presented here. The I-RBAC model is tested in the end through the execution outcome, which indicates a linear run-time trend of the model despite the existence of a massive amount of permission allocations and many queries. This trend is shown to validate the model's efficiency.

Thakare A. et al. [11] created a model that helps individuals with the uniform application of policies, minimises inefficiency and fecklessness, and incorporates the implementation of priority-dependent resource access rights throughout different users in a large business. Using the example of healthcare as a use case, the merits of the suggested model are examined by contrasting it with frameworks that already exist as well as the one that Azure employs. The outcomes of the study demonstrate that the suggested model categorises the policy algorithm which is based on priority features and demonstrates that the suggested framework can deal with issues that typically arise when attempting to deal with massive challenging conditions in large organisations. This is demonstrated by the fact that the suggested framework incorporates the priority feature facility that is present in the current RBAC model.

K. Rajesh Rao et al. [12] introduce a novel PEKS technique for string search. This technique, in contrast to prior designs, is distinct from bi-linear mapping, and it is 97% more efficient than the PEKS for text search that Ray et al. presented in TrustCom 2017. Secondly, it adds role-based access control, to multi-user PEKS. This allows an unlimited group of individuals to search for and access encrypted files depending on the responsibilities they have. R-PEKS is the name that has been given to this integrated program. When compared to the PEKS method, R-PEKS can achieve an effectiveness of up to 90%. These formal security assertions are presented for the various components of R-PEKS, and a commercial dataset is used to validate these techniques.

PreetiGoyal et al. [13] The concept of "cloud computing" is becoming increasingly well-known in the world of computer science. Cloud computing is a method that launches services in the form of storing and retrieving data via the internet rather than using the hard drives of machines. IaaS, PaaS, and SaaS are just a few of the many services that may be obtained through the cloud. Intruders have begun to disrupt the procedure to obtain confidential files belonging to a variety of cloud customers as the rise of cloud computing continues to rise. A system that can provide the essential characteristics for security needs to exist. Many different security principles, including secrecy, access control, integrity, and reliability, are implemented by cloud services to ensure users' safety. In the work that is being presented here, all of these ideas are put into practice in an environment by employing algorithms such as ECC to improve data confidentiality. In this case, MD5 is used to ensure the data's integrity on the server side, while RBAC techniques are used to control who has access to the data. Therefore, the high-security level required for the cloud infrastructure is satisfied by the architecture that was proposed.

A unique unified access control approach for diverse multi-tenancy frameworks in the cloud system is presented by MeryemAyache et al. [14]. This paradigm, which is known as XBAC, can manage several kinds of access control regulations. The access regulations and request details are both represented by a learning automaton in the model that has been developed. A cloud broker is utilised to accomplish the integration of XBAC into the cloud environment. In addition to this, the study lays out a fundamental architecture that describes the connection between the primary elements that make up heterogeneous cloud systems. The preliminary findings of the simulation indicate that the suggested model satisfies the needs of the environment containing heterogeneous access models by prohibiting the utilization of the resources of multi-tenancy cloud computing by users who are not authorised to use them.

A cloud-based solution is presented by Ningning Cui et al. [15] for the online tracking of tool settings during end milling. The implemented application that integrates the IoT platform for the tracking of tool parameters in the cloud to the machine tool and visual system for the identification of cutting chip size is the most interesting aspect of this study. The optical system is responsible for the acquisition and transmission of signals relating the chip area to the IoT application. These signals are then utilised as an indicator for determining the conditions of the tool. Additionally, the novelty of the strategy that has been presented lies in the artificial intelligence that has been incorporated into the framework. This artificial intelligence supervises the circumstance of a tool by identifying the trend of the existing cutting force, and it provides protection for the tool from large strain by adjusting the processing parameters. The research has practical importance because it led to the development of a new system for rapidly assessing the state of tools. This system improves machining efficiency, ensures cost savings, lowers investment costs as a result of its use of a sensor with lower upfront costs, and enables distant process monitoring on portable devices. To determine whether or not the monitoring system would be feasible, machining tests were carried out. According to the findings, the system that was designed with an artificial neural network for the identification of cutting force patterns is capable of reliably detecting tool damage and halting the procedure within 35 milliseconds. This was achieved. This research reports that the classification accuracy using an ANN was 85.3% without any errors in the detection of tool breakage. This demonstrates that the approach is both successful and practicable.

According to Iman Saeed et al. [16], a large number of people all over the world use cloud-computing services. Cloud services companies like Amazon, Microsoft, and Google supply many of the cloud's services. The storage service is one of the available options. It is used by a significant number of users to store and retrieve their data. However, users will not utilise any service other than those offered by service providers who prioritise the safety and security of their data. This article analyses and contrasts the privacy and security offered by Amazon AWS S3 vs Microsoft Azure Blob. Its purpose is to conduct an evaluation of its privacy and security compliance programmes as well as to investigate how well its access control system safeguards the data of its customers. This elucidates the primary distinctions between the two, as well as the parallels between them, based on the prior arguments.

According to the findings of Muthunagai, Anitha's et al. [17], it is a tough operation to keep client data stored on a private storage device that is located physically nearby. Cloud storage can alleviate this issue by keeping the user's documents in a remote database. The user will then be capable of retrieving their data from the cloud-based platform regardless of where it occurs to be physically stored. In this research, the above-said security issues are resolved by first segregating the data that is sent by the client, and afterwards encrypting the fragmented data using the augmented attribute-based encryption strategy that has been proposed. These steps are taken in to tackle the vulnerability that was discovered. After that, the encrypted data is gathered in many different places. The client's transferred information is shielded from being interpreted by any unauthorised third parties due to the confidentiality of secret data that is stored at multiple focuses. In the end, with the assistance of a decoding technique, the knowledge that had been stashed away in a particular location was brought to light. After that, it minimises the quantity of system traffic that was generated as a result of retrieving the individual data that was transmitted from a separate site. This occurred while the system was attempting to retrieve the information from the client.

A valid particular approach to attribute-based access control (ABAC) is described in Gadouche et al[18]. The technique that is recommended for use is the Event-B technique. In order to develop the right prototype methodically, the research makes use of the prior appropriate verification that was performed in order to do so. The prototype exemplifies the many degrees of consideration that can be acquired by completing responsibilities associated with the refining process. A good number of the characteristics of ABAC are discussed at each of the levels of refinement, beginning with the level that is the most advanced and unique and going down to the one that is the most fundamental. These characteristics are recorded in the proofs themselves, namely inside the behaviourin particular. The procedure can be found on the websites of the organisations that are in charge of the administration of social insurance.

According to Chakraborty et al. [19] the ABAC framework mining issue predicts that attribute estimation for different compounds, such as customers as well as items in the structure, are offered, furthermore to the acceptance state, from which the ABAC technique ought to be located. Because this is a one-of-a-kind circumstance, the developers create the ABAC RuleSet Availability issue and come up with a computational in addition to an unusualness assessment for the response to that question. Later it presents the idea of ABAC RuleSet Infeasibility Enhancement together with a calculation for the solution to the problem. According to Majid Afshar et al. [20] an ABAC framework is advocated for use as a framework in the field of human services. It utilised the ABAC engine in order to carry out and sustain the implementation of social insurance programmes. With that, circumstances were handled involving emergencies by utilising this expertise. YingjieXue et al. [21] investigated an extraordinary attribute-based access control situation wherein multiple users with distinct attribute sets can work together just to acquire entrance authorization if the data holder enables their joint effort in the access process. In this situation, YingjieXue et al. explored a situation in which multiple users can accomplish it together just to acquire entrance permission. Particularly, the scenario centred on a situation where the users were attempting to connect to a database. In this scenario, the customers were seeking to retrieve the database. A system for attribute-based controlled synergistic access control is proposed wherein interpretive centre points in the access hierarchy might be chosen as decision-makers. According to the findings of the security assessment, their recommended technique is capable of preserving the data's secrecy and possesses some other important attributes that are necessary for ensuring its safety. The comprehensive analysis of the implementation of the plot demonstrates that the recommended plan is efficient in the sense that it is concerned with expenses related to storage and counting.

Viswanath, G., and Krishna, P. V., were the ones who came up with the method for transferring information or data into a cloud-based architecture [22]. The data were encoded with the assistance of this technique before being uploaded to the cloud storage service. In addition, the results of the simulation are available there, along with a total of 2630 KB of data that has been encrypted. It has used a collection of real-time medical information to analyze the effectiveness of the method. Li, H. et al. [23] developed a technique of image compression for data compression, encryption algorithm, and validation. This method may limit the extended subcontracting of the costly storing it away and the total amount. The researchers developed a completely new encryption approach with the support of a stochastic process in order to safeguard the confidentiality of photographic images.

On the other side, [24] cast doubt on the veracity of this method and proposed that the RS-IBE algorithm could benefit from the application of self-updatable encryption as a means of increasing its level of effectiveness. After the implementation of data security measures became feasible and, frankly, essential across the board for all cloud storage access, the next evaluation method that gained significance was the degree of accessibility.

In addition to other characteristics, data creation as well as the accessibility of data is usually regarded to be an attribute that enables in the data access control as well as information security and validity, thereby increasing the CPABE one independent parameter at a time [25]. This is because the data access control helps ensure that the data is secure and valid. This is because both the generation of data and its access are considered to be characteristics that permit dual control of information access and the verifiability of data security. After conducting in-depth research on many different encryption methods and access control systems, it became abundantly clear that very few of these have addressed trust factors as the major constituent, and even within those that have, the concept of cyclic flipping is rarely acknowledged. This became abundantly clear that very few of these have discussed trust factors as the basic element.

Wu et al. [26] developed a novel high-speed Cauchy method to generate an encoding calculation in consideration of the equipment execution system for decoding. This method uses a simpler calculation methodology to produce the same conclusions as Cauchy. Zhang et al. [27] presented Ca-Co as a practical Cauchy coding technique for data storage in distributed systems. It makes use of an XOR operation-enabled Cauchy matrix to generate a scheduling sequence. In contrast to current methods, the system offers a useful data storing strategy. It uses the XOR function to swiftly produce a Ca-Co matrix. A unique erasure decoding strategy is provided by Tang and Cai [28]. In the suitable stockpiling framework, erasure coding is favoured as a defence against capacity node disappointments. The data is initially divided into blocks of data, which are subsequently coded to produce blocks of encoding. The lost data blocks are recovered using the decoding technique. The decoding faults are acknowledged in light of the evolving translating change matrix system. Erasure coding increases the efficiency of interpretation while using less bandwidth during modernization. On the other side, the fictitious analysis shows how accurate the procedure is. The system of data restoration benefits from the erasure-coded data insertion techniques that are now available. The operator first replaces all inactive nodes with a single new node before attempting to restore data. In reality, the Node Replacement Process (NRP) can take hours or even days. The improved statistics are lost once more as a result of the lack of durability. In circulating record systems, a few techniques are used to preserve information dependability and accessibility.

Erasure Coding (EC) is a well-liked technique for improving space efficiency. In general, it has input/yield corruption elements and incalculable execution encoding and decoding. Kim [29] suggests a buffering and joining technique that treats numerous I/O needs that emerge during encoding as a single demand. Dispersing the disk input/yield loads produced during decoding, it provides four recovery possibilities. The allocated storage system uses erasure coding, a typically complicated technology, to safeguard against capacity node disk failure. It lessens repetition while enhancing resilience. The data is first divided into b pieces. The b blocks are subsequently changed into c blocks using the encoding method. The missed c b blocks are recovered by the decoding method. The repetitiveness settings (b, c, and d) of EC are often given varied redundancy configurations by clients of distributed storage. Its foundation is a flawless balancing act between execution and internal failure adaptation. This work helps to observe a very low-likelihood encoding method for a design with available configurations that yield superior results. In the cloud, EC codes are utilized with little data repetition. The accuracy of the Ca-Co approach is shown via a simple fictitious inquiry.

Bian et al. suggested the Optimal Weakly Secure Minimum Storage Regenerating (OWSPM-MSR) method [31]. It was tested on a Linux system with an Intel Core i5 CPU and 4GB RAM, and its implementation is based on Jerasure. The implementation used a block size of 1024 and a Galois field size of 28. They contrasted many elements with the old approach, including calculation time and storage overhead. This approach employs the RS coding algorithm based on a Cauchynetwork. Division and multiplication can be transformed into subtraction and addition problems over limited fields, and XOR tasks can identify them. They also calculated the time spent translating during data recovery. It takes a little bit longer than Product-Matrix Minimum Storage Regenerating in the worst-case scenario. (PM-MSR). As the number of data blocks (k) and parity blocks (p) rose, the interval between procedures grew longer. (m). The time required by the OWSPM-MSR approach exceeds the time required to encode the data since it takes longer to invert a matrix.

[32] A hybrid clustering approach termed HCMX to identify a clustering solution of several, dynamically changing XML documents. as opposed to treating each new document as a new version, the number of documents that were impacted. Utilize distance information during the preliminary clustering phase, with the alterations responsible for the production of the document version, as opposed to comparing all members of the clustering solution to find a new clustering solution after changes in the preliminary one. Homomorphic compression, which preserves the original structure of the document, was utilized to increase clustering speed and response time. The [33] Keeping track of the hardware inventory, including how many disks, RAM, and CPU are currently in use, across several server platforms (Windows and Linux). We will also keep track of a server's relative services and the number of applications that are running on it. Highlight the servers' maximum disk, memory, and CPU consumption. It will offer visual alert configurations to help the administrator rapidly identify problems and their root causes. One advantage of the system is that you can manage the activity of N servers

in a little amount of time or all at once. Therefore, both the downtime and the organizations' QoS (quality of services) will grow. Any resource failure will be quickly identified. A single human resource may handle the entire N number of servers, making troubleshooting simple. Using a straightforward website mining technique, [34] product information from the pages of an e-commerce website is extracted. Based on the entropy value at each node in the HTML tag tree of the retrieved web page, it first identifies the set of product descriptions. For greater accuracy in product extraction, a set of association rules based on heuristic features is then used. The [35] creating a blockchain-based spectrum selling and sharing system that takes user privacy seriously and addresses security issues. Using a Distributed Blockchain Consortium System, mobile network providers can exchange spectrum without the aid of a third party. (DBCS). Data is well shielded from plaintext, stalker, and ciphertext assaults as a result, making it simpler to monitor, manage, and regulate services. Experimental findings demonstrated that the suggested DBCS performed and secured better than other traditional systems.

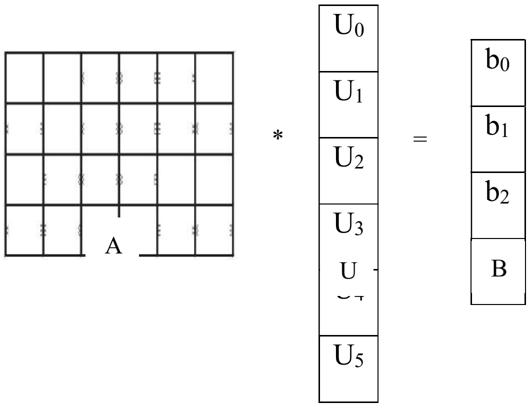
III. RESEARCH BACKGROUND

Cauchy Reed-Solomon Coding

Reed-Solomon (RS) codes are founded on a finite field, which is also sometimes referred to as a Galois field. When encrypting data with RS codes, it takes a lot of calculations to accomplish a Galois field mathematical operations (such as addition or multiplication), hence the performance is frequently inadequate. CRS codes are a modification of RS codes that offer two enhancements. To begin, a Cauchy matrix is utilised for CRS codes rather than a Vandermonde matrix. Furthermore, the Galois field multiplications that are used in CRS codes are converted into XOR computations.

The generation of Cauchy matrices is the most important aspect of CRS coding, and this is accomplished in the following manner. Considering a redundancy configuration of (k, m, r) in which $p + q \leq 2^r$, let U = u1,..., uq, V = v1,..., vp, and $U \cap V = \emptyset$, so that every ui and vj is a different element of $GF(2^r)$. After that, the Cauchy matrix in element (i, j) is calculated by using 1/(ui + vj) (where the addition, as well as division, are specified over the Galois field). Because the constituents of $GF(2^r)$ are the integers ranging from 0 to $2^r - 1$, each element e can be expressed by just a w-bit column vector, Ve (e), by employing the primitive polynomial across Galois Field [36]. In addition, every individual element e of $GF(2^r)$ can be transformed into the binary matrix Q(e), which has the dimensions of (r × r) and whose i-th column (i = 1,..., r) is equivalent to the column vector Ve $(e2^{i-1})$. Therefore, depending on the value of w, the Cauchy matrix can be transformed into a binary matrix of the form (qr*pr), which will be denoted as A.

Each data block, U, and each block of erasure codes, B, are each divided into w trips. In this manner, when there is a value of "1" present in each row of A, it is possible to perform XOR operations on the information that corresponds to those rows in U to retrieve the elements that makeup B. As can be seen in Figure 1, the erasure codes demand a total of 11 XOR operations.



$b_0 = U_2+U_3+U_4$

$b_1 = U_0+U_1+U_2+U_3+U_4+U_5$

$b_2 = U_1+U_2+U_3$

$b_3 = U_0+U_1+U_4+U_5$

Figure 1: Erasure Coding with Cauchy Reed-Solomon codes

When data are stored in a storage system that makes use of erasure codes, the data are encoded to achieve data redundancy. As a result, to increase the overall effectiveness of a system, it ought to cut down on the expense of erasure coding, also known as the amount of XOR operations. To accomplish this objective, there are two different encoding methodologies.

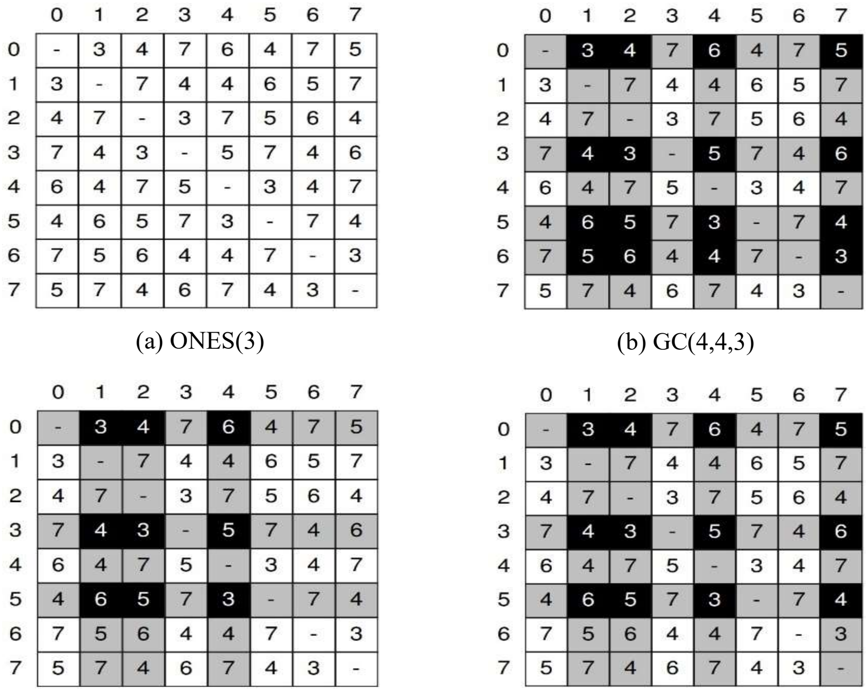
• The direct use of Cauchy matrices in the encoding of data. As can be seen in Figure 1, the concentration of a Cauchy matrix is what determines the amount of XOR operations as well as the efficiency of the encoding.

• Coding the information using the schedule. The encoding performance is determined by the order in which XOR operations are performed in the schedule.

Construction of Cauchy Matrices

When doing CRS coding with a binary Cauchy matrix, the amount of XOR operations required is proportional to the number of ones included in the matrix. Therefore, to obtain better efficiency, the bi-Cauchy matrix should include as few instances of the value one as is practically possible. Enumerating all of the possible Cauchy matrices is the simplest technique to get the best matrix among the many possible options. It is possible to generate Cauchy matrices, each of which may be utilised for encoding, if there is a given redundancy configuration with the parameters p, q, and r. As a result, the enumeration technique can only be utilised in situations in which the values of p, q, and r are quite low. If this is not avoided, the amount of time required to enumerate all of the matrices will be unacceptably long due to the combinatorial nature of the problem posed by the number of Cauchy matrices. Although certain methods, such as Original Cauchy, and Optimized Cauchy can produce a good matrix with fewer ones for greater w, this matrix might not be the best possible one[36].

- The Optimizing Cauchy heuristic first builds a ($2^r \times 2^r$) matrix, which is denoted as ONES(r). The element (i, j) of this matrix relates to the number of ones within binary matrix $Q(1/(i + j))$, which is the first step in the process of constructing a Cauchy matrix with the notation GC(p, q, r). The ONE's matrix is depicted in Figure 2(a). Following this, two distinct sets are chosen denoted $U = \{u_1, . . . , u_m\}$ and $V = \{v_1, . . . , v_k\}$from the range $\{0, 1, . . . , 2^r - 1\}$, in the manner described below.



	0	1	2	3	4	5	6	7
0	-	3	4	7	6	4	7	5
1	3	-	7	4	4	6	5	7
2	4	7	-	3	7	5	6	4
3	7	4	3	-	5	7	4	6
4	6	4	7	5	-	3	4	7
5	4	6	5	7	3	-	7	4
6	7	5	6	4	4	7	-	3
7	5	7	4	6	7	4	3	-

(a) ONES(3)

	0	1	2	3	4	5	6	7
0	-	3	4	7	6	4	7	5
1	3	-	7	4	4	6	5	7
2	4	7	-	3	7	5	6	4
3	7	4	3	-	5	7	4	6
4	6	4	7	5	-	3	4	7
5	4	6	5	7	3	-	7	4
6	7	5	6	4	4	7	-	3
7	5	7	4	6	7	4	3	-

(b) GC(4,4,3)

	0	1	2	3	4	5	6	7
0	-	3	4	7	6	4	7	5
1	3	-	7	4	4	6	5	7
2	4	7	-	3	7	5	6	4
3	7	4	3	-	5	7	4	6
4	6	4	7	5	-	3	4	7
5	4	6	5	7	3	-	7	4
6	7	5	6	4	4	7	-	3
7	5	7	4	6	7	4	3	-

(c) GC(3,3,3)

	0	1	2	3	4	5	6	7
0	-	3	4	7	6	4	7	5
1	3	-	7	4	4	6	5	7
2	4	7	-	3	7	5	6	4
3	7	4	3	-	5	7	4	6
4	6	4	7	5	-	3	4	7
5	4	6	5	7	3	-	7	4
6	7	5	6	4	4	7	-	3
7	5	7	4	6	7	4	3	-

(d) GC(4,3,3)

Figure 2: Example of creating Cauchy matrices using the Optimizing Cauchy heuristic

- If p = q and p is a power of two, then for p> 2, GC(p, p, r) includes the components of GC(p/2, p/2, w), and GC(2, 2, r) necessarily includes the column set Y = {1, 2}. This is because, for p > 2, GC(p, p, r) includes the components of GC(p/2, p/2, r). For instance, the expression GC(4, 4, 3) can be found in Figure 2 (b).
- When p equals q and p is not a power of two, GC(p, p, r) is defined by first creating G C(p′, p′, r), where p′ is greater than p and p′ is a power of two. This allows us to determine how to define GC(p, p, r). After that Next, superfluous rows and columns are deleted in alternating fashion until p by p matrix is obtained. Figure 2(c) demonstrates that to define GC(3, 3, 3), one row and one column from GC need to be removed (4, 4, 3).
- When p is not equal to q, GC(min(p, q), min(p, q), r) is constructed first, and then additional rows or columns added suitably. As can be seen in Figure 2 (d), GC(4, 3, 3) is built by including an additional column in GC (3, 3, 3).

To use the Cauchy Good heuristic, a Cauchy matrix is constructed which will refer to as GM. Then, split (defined over Galois field) each element of GM, like the one in column j, by GM0,j in such a way that GM is revised and the components of row 0 are all "1." In the remaining rows, like row i, the total number of ones is counted which is represented by the variable N. Then, the elements of row i are then splitby $GM_{i,j}$, and for each element of the row the number of ones is counted which is represented by the notation N_j ($j \in [0, p-1]$). The final step is to determine which value from the range $\{N, N_0, N_1 . . . , N_{k-1}\}$ is the smallest and then carry out the processes that produce it. Because of this, the Cauchy-Good heuristic allows to successfully generate a matrix.

The two methods discussed above have the potential to generate a binary matrix that has a lower number of ones; however, this matrix may not be the best option among the several

Cauchy matrices. Erasure coding is a procedure that uses XOR operations, and research has shown that there are lower limitations on the number of ones that may be contained within a Cauchy matrix [4]. This information was uncovered as part of the investigation into how to reduce the amount of XOR operations used. Consequently, the encoding performance cannot be significantly improved solely by decreasing the density of the Cauchy matrix. This presents a challenge.

Encoding with Scheduling

While carrying out data encoding with a particular Cauchy matrix, intermediate results could be made useful to cut down on the number of calculations that need to be performed twice.

Because of this, strategically timing these XOR operations might result in savings in CPU usage due to the reduced number of XORs that are carried out. A straightforward scheduling illustration is presented in Figure 3. After the schedule S of matrix A has been established, the generation among all erasure code elements needs only 6 XOR operations as opposed to the fewer than 11 times that are necessary when encoding utilizing Cauchy matrices explicitly. The concept of scheduling causes the times of XOR operations to exceed the barrier of the count of ones in the Cauchy matrix [36], which results in a reduction in the durations of XOR operations and an acceleration of subsequent erasure coding.

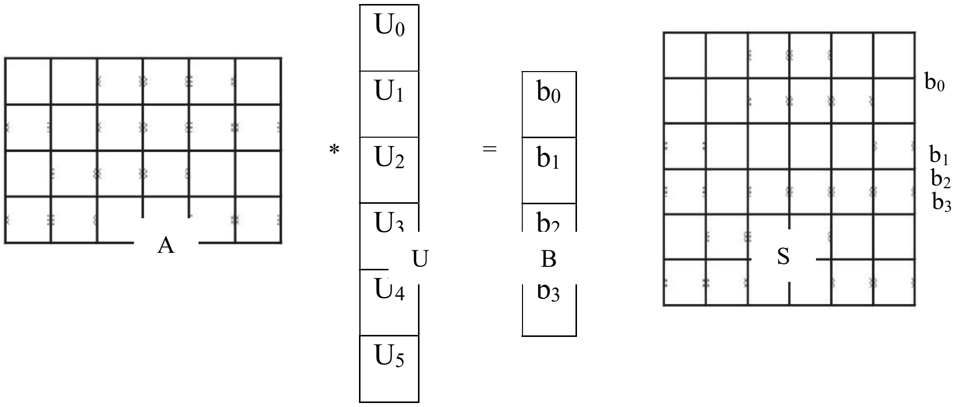


Figure 3: Erasure Coding with the best possible schedule

The following are some of the scheduling strategies that can be used to a Cauchy matrix.

- CSHR : to avoid doing the same computations several times, it employs the erasure codes that were previously generated to compute successive erasure coding elements.
- Uber-CSHR: this kind of CSHR represents an advancement. Not only does it make use of the elements of erasure coding, but it also makes use of the intermediate findings acquired to speed up subsequent calculations.
- X-Sets [36]: this algorithm recognizes and uses the typical XOR operations when it generates erasure codes to prevent performing the same computations several times. Calculating a particular XOR operation and storing the outcome in a way that allows later calculations to make direct use of it is something that can be done. There are a variety of methods available like MW, MW-SS, MW-Matching, MW2, Uber-XSet, and Subex, that can be used to pick exactly whichever common XOR operation comes into the schedule initially.

Creating a schedule out of a Cauchy matrix that reduces the total number of XOR activities as much as possible is still an outstanding challenge. The techniques that are currently in use for seeking schedules, like CSHR, Uber-CSHR, or X-Sets, are all heuristics, which means that it has no way to ensure that it will find the best solution. The problem of creating an optimal schedule using common sums was hypothesised to be an NP-Complete one.

IV. RESEARCH METHODOLOGY

To attain fault tolerance, this innovative strategy makes use of the Advanced Cauchy Reed Solomon method. The application of this concept involves listing down every conceivable arrangement of the data-matrix format and various schedulers, then selecting the most optimum technique from among those available. Because there is no specific rule or procedure that is defined to locate the optimal combination. The process of this is illustrated below:

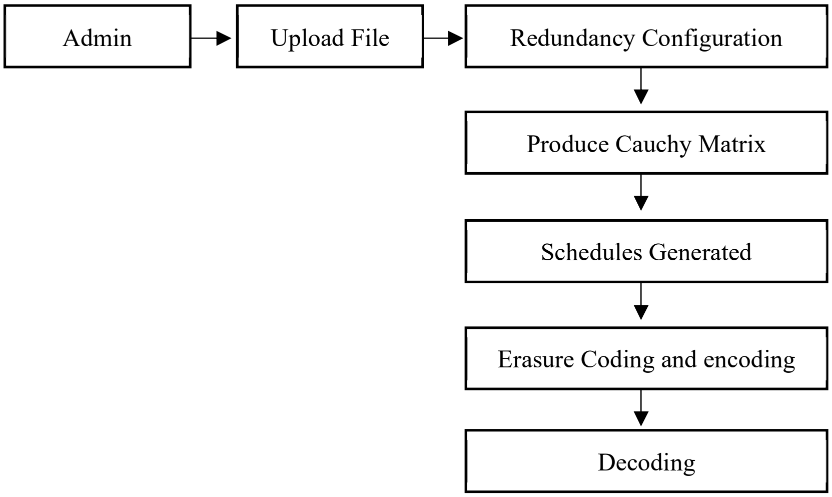


Figure 4: Framework of Proposed ACRS Technique

Becausethe ‘redundancy configuration' requirements vary depending on the user data, it is impossible to draw any conclusions. In addition, the decision for the same kind of "matrix computation as well as redundancy configuration" can be made by an individual based on the findings of an investigation.

As can be seen in figure 4, this discussion will focus on three primary elements, as well as the connections between them. The Cloudserver is located in the middle of the data owner as well as the data user. Both the data owner as well as the data user possess access privileges to the data stored on the cloud server, allowing them to upload and obtain the information. When a user uploads a file, they should be required to submit their login credentials so that the file can be authenticated. Users can upload files once they have successfully logged in. After that, the encoding of the file including preprocessing it is carried out. Both the references (indexes) to the first file that was generated and the actual file itself are saved. When a user uploads a subsequent file, the contents of that file are checked with the references included in the previously uploaded files. To preserve memory and cut down on the amount of time spent searching

When it comes time to save the information, it will generate a [8 * 8] matrix for every algorithm. Every data chunk is saved on the data node as well as matrix is placed into master mode. When a failure of any kind happens in the system, the data can be recovered with the assistance of the matrix and any rest of the data nodes.information, duplicate material is deleted and references are passed. The amount of time that this procedure takes up on the CPU can be seen here.

XOR Technique

Erasure coding can be implemented in its most basic form using the XOR technique. Let's say X, Y, and Z are the data cells. Therefore the parity cell is the XOR of these 3 data cells, which is written as x Ł y Ł z. When performing an XOR operation, just one parity bit is produced, and when any bit is lost, it may be restored using the rest of the data cells as well as a parity bit. Because it only yields one parity bit, the XOR operation can only withstand one failure for every n-sized group. This severely limits its usefulness. When the group size is n, the XOR operation has a fault tolerance of 1, and its storage efficiency is n-1/n.

Algorithm:

1. The system will initially calculate the values for (k,m,w), where k denotes the overall number of chunks, M denotes the matrix nodes, and w is the XOR of k and m.
2. The file is uploaded once all of the k, m, and w vectors have been followed.
3. The system will generate 4 chunks and employ data nodes to store the encrypted data when k = 4, m = 2, and w = 6. It will create a [8 * 8] matrix for each algorithm when it's time to save the data.
4. Each data chunk is saved on a data node, and the matrix is switched to master mode.
5. The matrix and any remaining data nodes can help retrieve the data using encoding when a failure of any kind occurs in the system.

Deduplication:

The deduplication of files should be removed from cloud storage to achieve the goals of saving storage space and enhancing the effectiveness of the cloud.

V. EXPERIMENTAL RESULTS

With the use of Cauchy matrices as well as XOR schedules that were generated by CaCo, an analysis was conducted to determine how well data can be encoded and decoded using the cloud storage system.

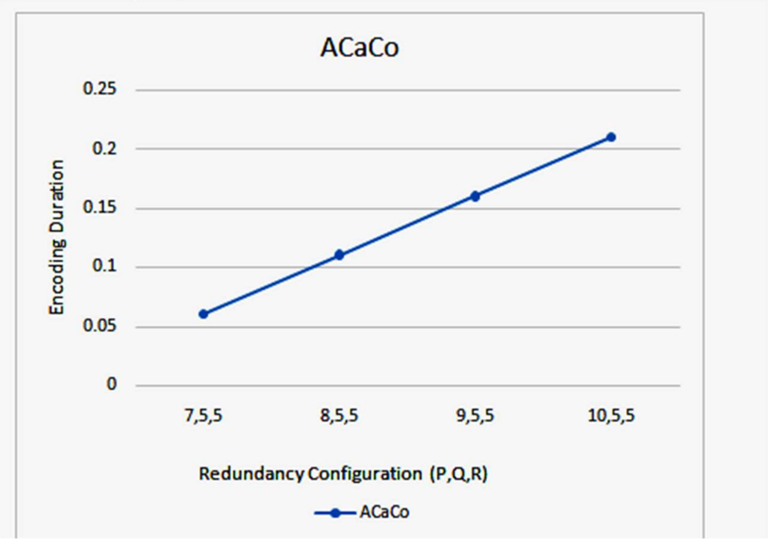


Figure 5: Encoding Duration of ACaCo with different Redundancy Configurations

The redundancy configuration (p, q, r) is changed by keeping q and r at 5, but raising p from 7 to 10. The data is encoded with ACaCo to provide parity data for a particular redundancy configuration, and the time required for the encoding process is recorded. Figure 5 presents a representation of the encoding time for ACaCo. First, as can be seen in the picture, the encoding times of CaCo exhibit an increasing trend along the x-axis as the parameter p is increased.

Table 1: Encoding Duration of ACaCo with different Redundancy Configurations

Redundancy	**ACaCo Encoding Duration**
7,5,5	0.06
8,5,5	0.11
9,5,5	0.16
10,5,5	0.21

As can be observed in table 1, the encoding time increases from 60 milliseconds to 210 milliseconds when ACaCo is utilised in the cloud system. The reasoning for this phenomena is that an increase in the value of p results in an increase in the number of data blocks that are encoded, which in turn results in an increase in the number of XOR operations.

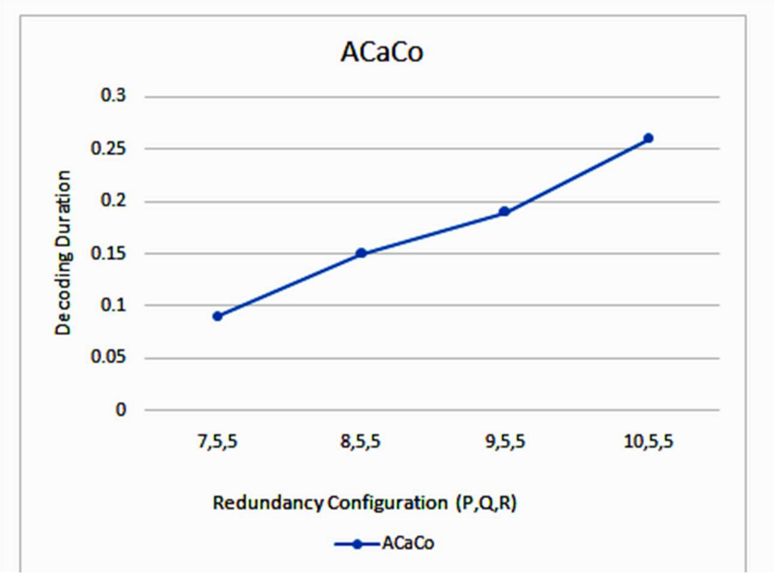


Figure 6: Decoding Duration of ACaCo with different Redundancy Configurations

The redundancy configuration (p, q, r) is changed by keeping q and r at 5 but raising p from 7 to 10. The data is decoded to reproduce any lost data with CaCo for a certain redundancy configuration, and the amount of time required for the decoding process is recorded. Failures of many discs at the same time can be tolerated thanks to the redundancy provided by these arrangements. When recovering from several instances of data corruption within the same group of p and q blocks, the computations required to do so have varying degrees of complexity. When ACaCo is first put to the test, it is performing what it accomplishes best: recovering from the failure of a single disk. The decoding timings utilizingACaCo are plotted in Figure 6, which depicts the process of recovering from one data disc failure. To begin, an increasing trend can be seen in the decoding timings of ACaCo along the x-axis as the value of p increases.

Table 2: Encoding Duration of ACaCo with different Redundancy Configurations

Redundancy	ACaCoDecoding Duration
7,5,5	0.09
8,5,5	0.15
9,5,5	0.19
10,5,5	0.26

As can be seen in table 2, the decoding time increases significantly when CaCo is utilised in the cloud system. It goes from 90 ms to 260 ms. The reason for this phenomenon is that a higher value for p indicates that a data decoding operation will require a greater number of data blocks and, as a result, a greater number of XOR operations. Below, figure 7 describes the time required in seconds for data encryption as well as decryption. Based on this experiment, the decryption could take high time than the encryption process.

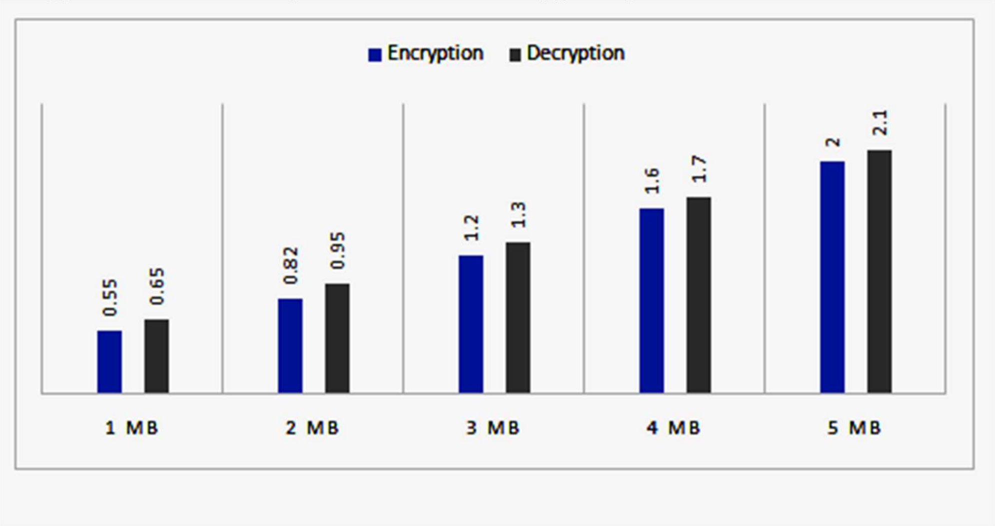


Figure 7: Time required in seconds for data encryption and decryption

The two-way encryption techniques are also carried out to achieve security to data during transmission and dynamic decryption at the selection of the destination server. In the below table, we demonstrate the complexity of the proposed and existing systems.

V. CONCLUSION

In this paper, after researching the various causes of data loss and methods for data regeneration via erasure codes, it is observed that increasing the count of 'XOR operations' improved the efficiency of the process. Prior studies do have a few drawbacks. As a result, to enhance fault tolerance for actual application data by discovering a novel method to recalculate loss data with the Cauchy Reed Solomon procedure. The data encryption and decryption approach can sometimes generate extra overhead due to the complexity of the algorithm. Thus, the dependency on multi-key algorithms also develops high computation. In the first experiment, we evaluated the system with five servers, and its required network utilization cost is around 1MB. Finally, the conclusion has arrived that reducing the number of XOR operations is not the only method that can be used to improve the performance of a deletion code. It's possible that other aspects of the code, such as the amount of data needed for recovery and the number of corrupted reads, could be a bigger performance bottleneck than the CPU overhead.

REFERENCES

1. Soni, Kritika, and Suresh Kumar. "Comparison of RBAC and ABAC Security Models for Private Cloud." 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). IEEE, 2019.
2. Rajput, Amitesh Singh, and Balasubramanian Raman. "Privacy-Preserving Smart Surveillance Using Local Color Correction and Optimized ElGamal Cryptosystem over Cloud." 2019 IEEE 12th International Conference on Cloud Computing (CLOUD). IEEE, 2019.
3. Hussein, Nehad H. "Cloud-Based Efficient and Secure Scheme for Medical Images Storage and Sharing using ECC and SHA-3." 2019 2nd Scientific Conference of Computer Sciences (SCCS). IEEE, 2019.
4. Jeong, Junho, et al. "Secure Cloud Storage Service Using Bloom Filters for the Internet of Things." IEEE Access 7 (2019): 60897-60907.
5. Cui, Bo, Zhikun Lan, and Xiangyu Bai. "Research on Role-based Access Control in IPv6 Smart Home." 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC). IEEE, 2019.
6. Sukmana, Muhammad IH, et al. "Unified Cloud Access Control Model for Cloud Storage Broker." 2019 International Conference on Information Networking (ICOIN). IEEE, 2019.
7. Yang M, Huang M. An Microservices-Based OpenStack Monitoring Tool. 10th International Conference on Software Engineering 2019 Oct 18 (pp. 706-709).
8. Soni K, Kumar S. Comparison of RBAC and ABAC Security Models for Private Cloud. In2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) 2019 Feb 14 (pp. 584-587). IEEE.
9. Ghafoorian M, Abbasinezhad-Mood D, Shakeri H. A thorough trust and reputation-based RBAC model for secure data storage in the cloud. IEEE Transactions on Parallel and Distributed Systems. 2018 Sep 17;30(4):778-88.
10. Ghazal R, Malik AK, Qadeer N, Raza B, Shahid AR, Alquhayz H. Intelligent Role-Based Access Control Model and Framework Using Semantic Business Roles in Multi-Domain Environments. IEEE Access. 2020 Jan 9;8:12253-67.
11. Thakare A, Lee E, Kumar A, Nikam VB, Kim YG. PARBAC: Priority-Attribute-Based RBAC Model for Azure IoT Cloud. IEEE Internet of Things Journal. 2020 Jan 3;7(4):2890-900.
12. Rao KR, Ray IG, Asif W, Nayak A, Rajarajan M. R-PEKS: RBAC Enabled PEKS for Secure Access of Cloud Data. IEEE Access. 2019 Sep 16;7:133274-89.
13. Goyal P, Makwana H, Karankar N. MD5 and ECC Encryption based framework for Cloud Computing Services. In2019 Third International Conference on Inventive Systems and Control (ICISC) 2019 Jan 10 (pp. 195-200). IEEE.
14. Ayache M, Gawanmeh A, Al-Karaki JN. XBAC: A Unified Access Control Model for Heterogeneous Multi-Tenancy Cloud Environments. In2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC) 2019 Jun 24 (pp. 1872-1878). IEEE.
15. UrošŽuperl, Krzysztof Stepien, Goran Munđar and MihaKovačič A Cloud-Based System for the Optical Monitoring of Tool Conditions during Milling through the Detection of Chip Surface Size and Identification of Cutting Force Trends, 2019 IEEE.
16. Saeed I, Baras S, Hajjdiab H. Security and Privacy of AWS S3 and Azure Blob Storage Services. In2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS) 2019 Feb 23 (pp. 388-394). IEEE.

17. S. U. Muthunagai and R. Anitha, “Secure Access Control Method in Cloud Environment Using Improved Attribute Based Encryption Technique,” Int. J. Engineer. Adv. Tech., 2019.
18. H. Gadouche, Z. Farah and A. Tari, “A correct-by-construction model for attribute-based access control,” Clust. Comp., pp. 1-12, 2019.
19. S. Chakraborty, R. Sandhu and R. Krishnan, “On the feasibility of attribute-based access control policy mining,” in IEEE 20th Int. Conf. Inf. Reuse Integrat. Data Sci. (IRI), July 2019, pp. 245-252.
20. M. Afshar, S. Samet and T. Hu, “An attribute based access control framework for healthcare system,” J. Phy. Conf. Ser. vol. 933, 2018, p. 012020.
21. Y. Xue, K. Xue, N. Gai, J. Hong, D. S. Wei and P. Hong, “An attribute-based controlled collaborative access control scheme for public cloud storage,” IEEE Tran. Inf. Forens. Security, vol. 14, pp. 2927-2942, 2019.
22. Viswanath, G., & Krishna, P. V. (2020). Hybrid encryption framework for securing big data storage in multi-cloud environment. Evolutionary Intelligence, 1-8.
23. Li, H., Yu, C., & Wang, X. (2020). A novel 1D chaotic system for image encryption, authentication and compression in cloud. Multimedia Tools and Applications, 1-38.
24. K. Lee, “Comments on secure data sharing in cloud computing using revocable-storage identity-based encryption,” IEEE Transactions on Cloud Computing, vol. 8, no. 4, pp. 1299–1300, 2020.
25. Q. Zhang, S. Wang, D. Zhang, J. Wang and Y. Zhang, “Time and attribute-based dual access control and data integrity verifiable scheme in cloud computing applications,” IEEE Access, vol. 7, pp. 137594– 137607, 2019.
26. R. Wu, L. Wang, and Y. Wu, “A High-Speed Cauchy CODEC Algorithm for Distributed Storage System”, In Proceedings of the 2020 International Conference on Internet Computing for Science and Engineering (ICICSE '20). Association for Computing Machinery, New York,NY,USA,pp.20–24, 2020,
27. G. Zhang, G. Wu, S. Wang, J. Shu, W. Zheng and K. Li, "CaCo: An Efficient Cauchy Coding Approach for Cloud Storage Systems," in IEEE Transactions on Computers, vol. 65,no.2,pp.435-447,1Feb.2016,doi: 10.1109/TC.2015.2428701.
28. D. Tang and H. Cai, "A Novel Decoding Method for the Erasure Codes", Security and Communication Networks, vol. 2021, Article ID 8755697, pp. 1-12, 2021,.
29. J-J Kim, “Erasure-Coding-Based Storage and Recovery for Distributed Exascale Storage Systems”, Applied Sciences, Vol. 11, No. 8:3298, https://doi.org/10.3390/app11083298.
30. X. Wang, Z. Zhang, L. Shan, J. Li, Y. Wang, H. Cao, Z. Li, “An Optimized Encoding Algorithm for Systematic Polar Codes”, J Wireless Com Network 2019, 193 (2019),pp.1-12,

    J. Bian, S. Luo, Z. Li and Y. Yang, "Optimal Weakly Secure Minimum Storage Regenerating Codes Scheme," in IEEE Access, vol. 7, pp. 151120-151130, 2019,
31. Sonawane, V. R; Rao, D. R. (2015). HCMX: An Efficient Hybrid Clustering Approach For Multi-Version Xml Documents. Journal of Theoretical and Applied Information Technology, 82(1), 137.
32. Sonawane, V. R., Singh, L. L., Nunse, P. R.,Nalage, S. D. (2015, December). Visual monitoring system using simple network management protocol. In 2015 International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 197-200).
33. Sonawane, V. R.,Halkarnikar, P. P. Web Site Mining Using Entropy Estimation. In 2010 International Conference on Data Storage and Data Engineering.

34. Liu, L., Shafiq, M., Sonawane, V. R., Murthy, M. Y. B., Reddy, P. C. S., kumar Reddy, K. C. (2022). Spectrum trading and sharing in unmanned aerial vehicles based on distributed blockchain consortium system. Computers and Electrical Engineering, 103, 108255.
35. Guangyan Zhang, Guiyong Wu, Shupeng Wang, Jiwu Shu, Weimin Zheng, and Keqin Li, “CaCo: An Efficient Cauchy Coding Approach for Cloud Storage Systems,” IEEE Transactions on computers, Vol.62, No.11, November 2015.