ISSN NO: 2770-4491
Date of Publication:28-06-2022

# Cloud Computing Services with Minimal Redundancy, Storage, and Effective Processing Capabilities

**Salam Abdulabbas Ghanim Ali Al_Hachemi**
The Republic of IRAQ, Ministry of Education,
Directorate - General of Dhi QAR, Department of Education REFAEE
salamalghanim@gmail.com

**Abstract:** To put it another way, cloud computing is a method that allows for the rapid movement of resources such as servers, storage, networks and services from one location to another utilizing little administration effort. For the most part, it is a large-scale distributed computing architecture that is based on dynamically-scalable and managed virtualized processing power, storage, applications and services. The cloud presents to users as a single point of access for all of their anticipated computing requirements. Customers may keep their data in the cloud and use the on-demand software from the cloud without having to worry about local infrastructure. When it comes to computers, security and reliability are two of the most pressing concerns. A revolutionary new platform has been developed to solve the issues of security vulnerability and data center dependability in this study. The Cloud Storage system employs data redundancy strategies including replication, RAID, and erasure code to give tools at any time (reliability). Cloud computing systems may be installed in one of three ways. Public, private, and even a mix of all three have existed. There is already a private cloud at which the cloud is developed, managed by the particular enterprise in the proposed system. Replication and RAID's drawbacks were mitigated by the inclusion of Erase Code in the private cloud system's design. Low storage costs, normalization costs, fast processing speeds, ease of data center recovery in the event of a disaster, and assistance in setting up a secure cloud storage system are just some of the benefits it provides.

**Keywords:** Data replication, RAID, Erase code, computing power, and storage cost.

## I. Introduction

It creates a large number of copies of this object to store. The data center has a lot of storage space. Considering that all data may be provided locally, Replication also provides better or more indirect virtual server performance. When a replication system is used in the storage system, the machine's durability, dependability, and accessibility are all improved, as are the speed and accuracy of queries and answers. A centralized solution, on the other hand, needs greater storage space. The cost of updating is expensive, and ensuring data integrity across various security components is difficult because of the huge normalization throw. On-demand, on-demand system entrance into your common pool of customizable computing tools (e.g., server, storage, systems and services), which may thus be instantaneously created and transferred with a modest direction campaign [3] is made possible by cloud computing platforms. Customers may keep their data in the cloud and acquire the remote-requirement call-loud software they need without any infrastructure limitations. The cloud storage platform makes use of data partitioning techniques like Replication and Erasure Code to make the tools readily available at any time.

Traditional encryption methods such as DES or AES may be used to encode the message using either Reed-Solomon code (RS) or even Tornado before preserving the information, thereby reducing storage disc space and cloud data center procurement. Encryption and decryption are performed using a symmetric-key, whose primary is protected from virtual servers. After the data owner's permission, users may use the key to decrypt the information. An overview of the various approaches and how they favor shorter storage distances is provided in this chapter. The most significant contribution of the job is the calculation and formation of various storage sizes and calculation durations of the combination of multiple erasure codes with varying safety calculations. In comparison to replication, erasure codes are more cost-effective, require less storage space, are more durable, and are easier to retrieve. Each block of k message emblems is encoded straight to some code sentence symbols, with the addition of n-k test symbols generated from the message emblems. The n-character code is kept in a different data center. Using Reed-Solomon code and erasure code, this chapter aims to make it easier to store data in cloud data centers

## II. Literature Review

Peter Kunsz and his colleagues (2005) Give examples of the early data storage schemes that used data replication to keep a big power-spread database at high availability, fault tolerance, and extremely low access time. Data will be replicated and stored in many data centers in order to reduce the time it takes to retrieve information. In addition to regulating their period lengths, it is necessary to maintain consistency and up-to-datedness of these replication copies. Document transfer protocols and services (grid FTP), copies, and metadata catalogues have all been well taken care of via replica management. This grid FTP enables the user to do their own copy selection process for obtaining user information. Modules such as replication management service center components, optimization, subscription, balancing, session management, and security are all discussed in this article.

Cong Wang et al (2012) offered a supply storage ethical audit system without the need of server hardware or application involvement. ' Users may reorganize cloud storage operations using nominal and coded data using this design technique. As a result, the malfunctioning function cloud, as well as r storage correctness, have been discovered. Data corruption and host misbehavior may be rapidly identified by the Administrator during the information storage correctness confirmation, even if the same level of storage correctness confidence has been kept by using token computation and erasure-coded data. Dispersed 'm+k' servers have been set up to prevent the information file from collapsing. In this case, the authors inserted tokens on the vectors before the file supply. Additionally, the faulty data is reconstructed while these encrypted components are saved on various cloud servers.

They depend on pre-computed token confirmation to assure data storage accuracy and to pinpoint the location of data malfunctions. Before the provision of documents, a person estimates a certain number of tokens to be used. Each token consists of two blocks of data. It takes a user a long time to earn the storage and accuracy of the data because of all the blocks that cloud servers make while calculating and bringing difficulties to an individual. Individuals have the option of either storing tokens on their own or encrypting them and maintaining them on cloud-based servers.

Song tao Liang et al (2014) Recommendations for reducing the fixing bandwidth and disc I/O costs from the storage strategies were made by using hybrid storage re-generating codes. A repair theory may be categorized into three groups. There are three stages of repair: operational, accurate, and hybrid vehicle. The authors used just two items in their list of repair methods: exact restoration HMSR codes. Using a quick regeneration technique, these HMS Janin codes outperform ordered M SR codes in terms of data downloading. The HMSR algorithm completes rehabilitation by repairing the storage system under the small dimensions of the limited region.

In 2016, Yingxun Fuet al Only disc failures should be retrieved using stack-level retrieval, rather than strip-level recovery. Both the greedy algorithm and the rotational retrieval method are extensively used in this technique's two regaining y mechanics. Both a balance-priority strategy and a search-period priority approach are used by the greedy algorithm. This fine-grained policy-based access control and document deletion system was presented by Yang tang and others in 2012. It's nothing more than a safe and secure way to store data on the cloud. Fine-grained policy set access control and document deletion have been achieved with this method. Key managers are used to maintain FADE's in-built library of cryptographic keys. Attribute-based encryption is used to provide fine-grained access control, and the fault-tolerant main direction is used to ensure deletion of files. An automated system Employing cloud-based key management agencies, managers and companies must dedicate FADE to their employees. Working with a quorum technique, on the other hand, may help strengthen the network. The files were removed as the coverage reached page 139.

ISSN NO: 2770-4491
Date of Publication:28-06-2022

## III. Error correction methods

Only disc failures should be retrieved using stack-level retrieval, rather than strip-level recovery. Both the greedy algorithm and the rotational retrieval method are extensively used in this technique's two regaining y mechanics. Both a balance-priority strategy and a search-period priority approach are used by the greedy algorithm. This fine-grained policy-based access control and document deletion system was presented by Yang tang and others in 2012. It's nothing more than a safe and secure way to store data on the cloud. Fine-grained policy set access control and document deletion have been achieved with this method. Key managers are used to maintain FADE's in-built library of cryptographic keys. Attribute-based encryption is used to provide fine-grained access control, and the fault-tolerant main direction is used to ensure deletion of files. An automated system Employing cloud-based key management agencies, managers and companies must dedicate FADE to their employees. Working with a quorum technique, on the other hand, may help strengthen the network. The files were removed as the coverage reached page 139.

The Erasure coded cloud storage system divides N disc into M discs and K discs, with M discs containing n code word symbols and K discs containing n-k code information (check bits). The erasure code is used to get the check bits from the k message symbols and the co de word symbols. An erasure code has two distinct features. In the first place, it should be MDS (maximum Distance Separable), which means that if any K of the N discs are out of order, their information may be recomputed from an active M disc alive. It's the new disc (data center) that rebuilds whatever was on the failed disc. The M data discs also include raw data since erasure codes are systematic. In the MDS, the RS codes are well-known

| Operations | erasure code | Optimal erasure code |
|---|---|---|
| Time for encoding | $(k+ŧ)\ln (\frac{1}{e})^{p}$ | $(kŧ)^{p}$ |
| Time for decoding | $(k+ŧ)\ln (\frac{1}{e})^{p}$ | $(kŧ)^{p}$ |
| Inefficiency of decoding | 1 | 1+e |
| General calculation | polynomial | X-OR |

There is a Low-Density Parity Check (LDPC) chart that is part of tornado code, and all of the tornado codes have been made for reliable multicast. To create parity nodes, the data logos and parity symbols of each node have been kept. All tornado codes are derived from bipartite graphs that cascade. The number of steps and nodes required to produce a Tornado Code might vary widely. However, the fault tolerance qualities are owned by the specific border degree distribution that results in a cascaded bipartite LDPC chart. Erasure codes, as opposed to replication, provide a higher level of mistake tolerance. [8].

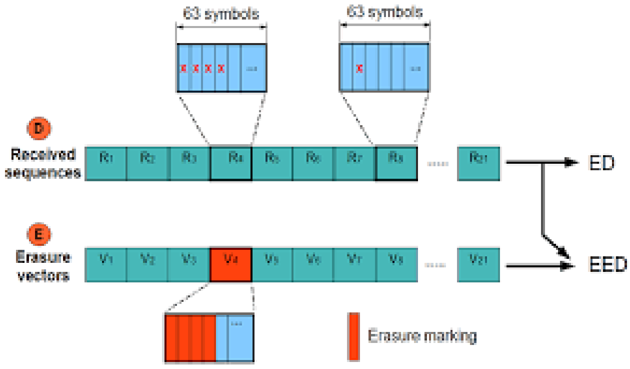

Figure 1: Function of encoding system

All procedures are carried out individually by the datacenters in order to function in a scattered environment. Adware, adware, and message routing are all necessary components of this strategy. It is possible for cloud storage systems to satisfy certain criteria of information security and data resilience and storage due to the integration of monitoring and partitioning operations

## IV. Model setup

SS inch, SS 2 are two of the n datacenters (storage servers) in this system. In the simulation, your investment's return is represented by eight data centers and four key managers. These data owners are protected by the symmetric secret that is managed by the important managers. Erasure's cloud-storage technology has three phases: data storage, data storage, and data recovery. The data recovery phase is the last phase. During the data storage time, data owners use any of their symmetric techniques to detach the communication and send it to data centers. A message M is broken up into M1, M2, M3, M4.... Mk (k=8) because it is encrypted into cypher text ci utilizing 8 data centers. SSi(i=8 storage server) with identification I D randomly selects data centers to receive these encrypted text blocks. It is saved at each data center after receiving encrypted messages from the data owners. k encoded message blocks cannot be received by any storage data center. Users may destroy received messages from symmetric-key (which is stored in primary Managers) with the permission of the information owner after the data shredding time by receiving decoded messages from information centers using identifier i-d. The administrator's petition is used to restore pieces of data that are critical to key supervisors and live data centers when the information center collapses

## V. Proposed method

Replicating data in real time may improve the performance of an application system. We are currently calculating the prevalence level and copy element in order to find the most suitable document to replicate and decide on the different replicas. To figure out where to put the copies, we're using fuzzy logic. We also use a round-robin method to deploy the replicas from the identified systems.

***Algorithm for Guaranteed Reliability***

```
1: $t \leftarrow GetTime()$
2: $A \leftarrow Proc - (H)Digitalize\ sequence$
3: $for\ tier \leftarrow Client\ tier - 1\ down\ to\ RootTier + 1\ do$
4:     $A \leftarrow Aggregate(A)$
5:     $for\ all\ record\ r€A\ do$
6:         $if\ r, numOfAccesses \geq threasholds$
7:             $\mathrm{Update - Ctime(r.fileID, r.nodeID, t)}$
8:             $\mathrm{Getreplicate\ (r.fileID, r.nodeID, t)}$
9:         end if
10:     end if
11:     end for
12: end for
```

Data from A is combined into the current grade in the first algorithm lineup 4. The specifics of Aggregate's work will be studied to a considerable extent using a large data set. All recordings' node IDs are derived from the current processing grade after the aggregate. It will process further for each and every album of RIN A if r.numOfAccess is more than the current tier's limit (line 6). r.fileID exists at the node of all r.nodeIDs in the event. This is followed by the current replication session duration being updated and ep being removed from A. Alternately and regardless of whether or not the r.nodeID node has sufficient distance for document r.fileID to replicate and expel record dtc from the node (lines 10—1 2 ). It's probable that the remaining recordings in A will be flashed into a better inline 4 when the internal loop is finished. As previously stated, the improved option A will also be handled. Calculating the ratio between your seldom sought file on the host and the full collection of often hunted files allows you to rank the optimal document. There is an order to the most often used files. Most of your user's replies are saved in the main cloud, which is where the files are often accessed. A significant amount of calculation and article processing may be required as a result of the host's new status. In addition, it is widely accepted that cloud computing systems often include a large number of servers. The mechanism used to determine rank makes it easy for anybody with a valid domain name to have access to cloud server data. On each server, an estimate of the number of files that can be found is computed. rnk calculates low-status documents based on the cloud host's prediction and the minimal status files. As a result, the file's reduced status chooses just the bare minimum of options. The next step is to choose a 'I' ranking of services that adheres to the predetermined shrewd preferences only insofar as a viable standing structure has been recognized. Replicated data is removed from the second host depending on the brink of a shutdown of the selected software and saved only on two servers. Additionally, the status values may be calculated and deleted from files based on RAM

## VI. Results

**Storage Size:** Simulator findings must be used to verify the functioning of the erasure codes platform. After shredding (preparing to save) from the information centers, the chunked file size hasn't improved substantially. As a result of this, there are minor differences across techniques. The plain-text cube's dimensions are the same as the ciphertext's size because of encryption.

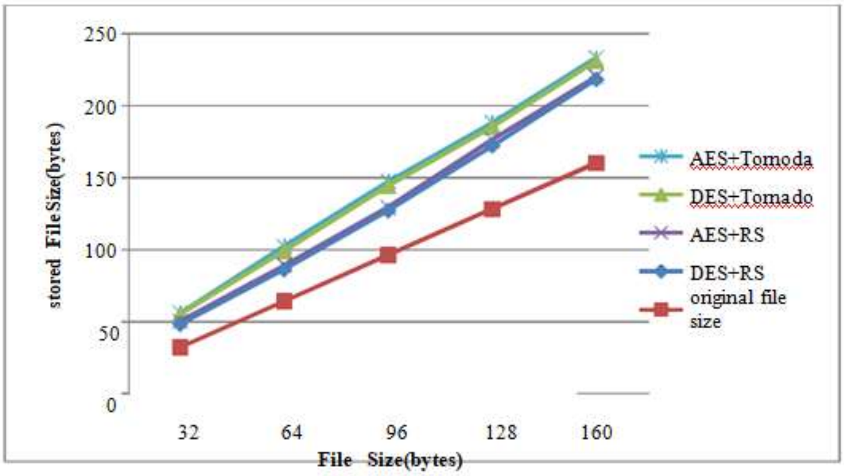

Figure 2: Comparison between original block and various encodes block size

Simulator findings must be used to verify the functioning of the erasure codes platform. After shredding (preparing to save) from the information centers, the chunked file size hasn't improved substantially. As a result of this, there are minor differences across techniques. The plain-text cube's dimensions are the same as the ciphertext's size because of encryption. Even if AES uses various keywords (such as 128,192,256), the distance between this ciphertext and the preceding block of plain text is not affected by the length of the key. The size of the ciphertext and the plaintext will be the same. It is also included in the chunks when the information size is expected to increase during the communication. In storage locations where data replication mechanisms are not used, this variation is mirrored. Techniques that replicate data need n data centers, but methods that utilize erasure codes only need n -k acceptable data centers, the results of each approach vary. RS codes take up less store space when compared to T-O tornado codes since RS codes analyses just a restricted number of elements throughout the document retrieval process. The tornado code, on the other hand, requires a vast quantity of data. The size of the charts is continually increasing as more data is added to the system, and as a result, more charts are parallel.

Reed Solomon codes, as predicted, perform better throughout the processing phase. The tornado code's performance is greater to that of the RS code, however the tornado co p is bigger and constructed utilizing probabilistic assumptions and indications. Because of this, Tornado code, on average, requires more data to recover a random file than Reed Solomon code. The table below provides information on the processing time required to save data from the data center for each of the many procedures.

Table 1: processing time comparison table

| **File size (Bytes)** | **DES and RS (ms)** | **DES and Tornado (ms)** | **AES and RS (ms)** | **AES and Tornado (ms)** |
|---|---|---|---|---|
| 32 | 0.731297 | 0.537067 | 0.621 283 | 0.48319 |
| 64 | 0.853517 | 0.603454 | 0.709 421 | 0.57473 |
| 96 | 1.135577 | 0.759629 | 0.905 823 | 0.71143 |
| 128 | 1.312048 | 0.879186 | 1.045 932 | 0.83172 |
| 160 | 1.628691 | 0.985965 | 1.168 921 | 0.89668 |

ISSN NO: 2770-4491
Date of Publication:28-06-2022

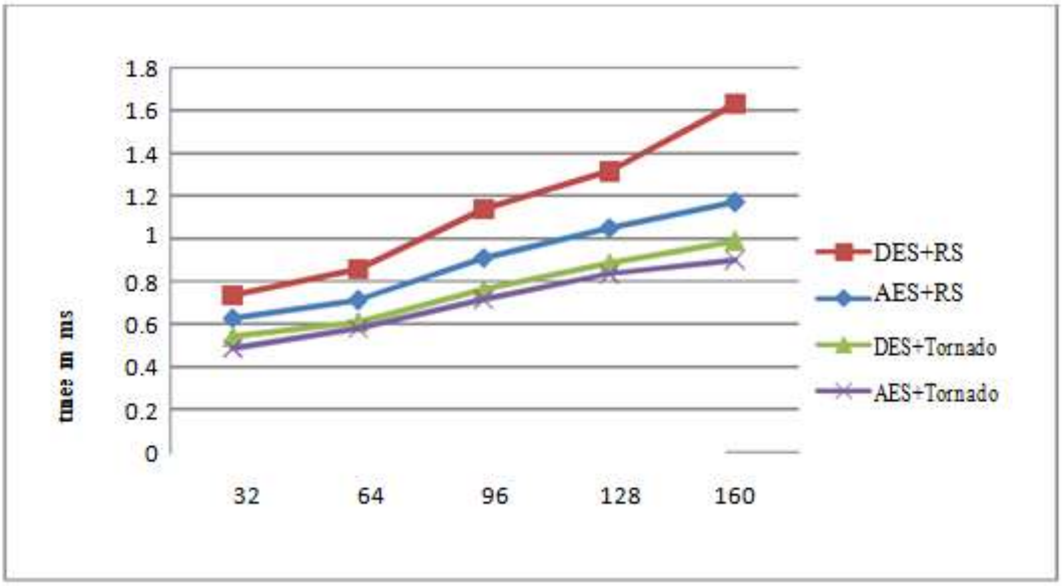

Figure 3: Comparison between various schemes storing times

Probabilistic and sophisticated graphs were utilized in R S code during the latest era of tornadoes. Reed Solomon codes, one of the four techniques of AES, deliver secured information with a low processing time.

## VII. Discussion And Conclusion

We've had great success with this technology, which is able to communicate and store data securely. The AES256 security implications, shared with its own, have certain programming libraries, and well-suited applications have made it one of the most widely used algorithms of its kind. While the procedure is easy to use, it requires a level of security for writing that is difficult to provide without a level of resilience that is impossible to achieve. As compared to the DES algorithm, the AES method treats 128 parts as one block while still being less responsive to striking and implementing quickly. In compared to replication, erasure coding consumes much less resources than does the latter. Using Reed-Sensible Solomon's Storage Space, and the R S code in the network, the application determines the time it takes AES to encrypt a protected chunked file and saves it to the information center.

## VIII. Future Scope

The focus of this essay was only on the recovery of a single data center from collapse. Community cloud security products and service companies might benefit from this paradigm. It's possible that T-study O's may have required more than one data-center retrieval and T-use Os of a variety of different types of software.

## References

[1]. HaiyingShen, Guoxin Liu, “Swarm Intelligence based File Replication and Consistency Maintenance in Structured P2P File Sharing Systems” IEEE Transactions on Computers, Vol. 64, No. 10, Oct 2015.

[2]. SameeUllah Khan, Ishfaq Ahmad “Comparison and analysis of ten static heuristics-based Internet data replication techniques” Parallel Distrib. Comput. 68 (2008)

[3]. Zheng Yan, Lifang Zhang, Wenxiu Ding, and QinghuaZheng, “Heterogeneous Data Storage Management with Deduplication in Cloud Computing” IEEE Transactions on Big Data, Vol. pp, No.99, May 2017

[4]. Jing Zhao,XuejunZhuo, “Contact Duration Aware Data Replication in DTNs with Licensed and Unlicensed Spectrum” IEEE Transactions On Mobile Computing, Vol. 15, No. 4, April 2016

[5]. Jenn-Wei Lin, Chien-Hung Chen “QoS-Aware Data Replication for Data Intensive Applications in Cloud Computing Systems” IEEE Transactions on Cloud Computing May 2014

[6]. Rodrigo N. Calheiros, Rajkumar Buyya “Meeting Deadlines of Scientific Workflows in Public Clouds with Tasks Replication” IEEE Transactions on Parallel and Distributed Systems, Vol. 25, No. 7, July 2014

[7]. Han Hu, Yonggang Wen, Tat-Seng Chua, Jian Huang, Wenwu Zhu and Xuelong Li "Joint Content Replication and Request Routing for Social Video Distribution over Cloud CDN: A Community Clustering Method" IEEE Transactions on Circuits and Systems for Video Technology, Vol. 26, No. 7, July 2016.
[8]. S.Annal Ezhil Selvi and Dr. R. Anbuselvi, "Ranking Algorithm Based on File's Accessing Frequency for Cloud Storage System", International Journal of Advanced Research Trends in Engineering and Technology (IJARTET) Vol. 4, Issue 9, Sep 2017.
[9]. Jonathan L. Krein, Lutz Prechelt "Multi-Site Joint Replication of a Design Patterns Experiment using Moderator Variables to Generalize across Contexts" IEEE Transactions On Software Engineering, Vol. X, No. X, Month 2015
[10]. Wenhao Li, Yun Yang, Dong Yuan, "Ensuring Cloud Data Reliability with Minimum Replication by Proactive Replica Checking", IEEE Trans. Computers 65(5): 1494-1506 (2016)
[11]. YaserMansouri, Adel NadjaranToosi, and Rajkumar Buyya "Cost Optimization for Dynamic Replication and Migration of Data in Cloud Data Centers" IEEE Transactions On Cloud Computing, Vol. pp, No. 99, January 2017
[12]. Runhui Li, Yuchong Hu, and Patrick P. C. Lee "Enabling Efficient and Reliable Transition from Replication to Erasure Coding for Clustered File Systems" IEEE Transactions On Parallel And Distributed Systems, Vol. pp, No. 99, March 2017.
[13]. Jerry Chou, Ting-Hsuan Lai "Exploiting Replication for Energy-Aware Scheduling in Disk Storage Systems" IEEE Transaction on Parallel and Distributed Systems, Volume 26, No 10, Oct 2015.
[14]. Guoxin Liu, HaiyingShen, Harrison Chandler "Selective Data replication for Online Social Networks with Distributed Datacenters" IEEE Transactions on Parallel and Distributed Systems, Vol. 24, No. 8, August 2016
[15]. Amina Mseddi, Mohammad Ali Salahuddin "On Optimizing Replica Migration in Distributed Cloud Storage Systems" 4th IEEE International Conference on Cloud Networking (IEEE CloudNet 2015)
[16]. Kan Yang &XiaohuaJia 2012, 'Data storage auditing service in c loud computing: challenges, methods and opportunities', Springer, World Wide Web, vol. 15, pp. 409-428.
[17]. Kan Yang & XiaohuaJia 2013, 'A n Efficient and Se cure Dynamic Audi ting Protocol for Data Storage in Cloud Computing', IEEE Transactions o n Parallel and Distributed Systems, vol. 24, n o. 9, pp. 1717-1726.
[18]. Rashmi, KV, Nihar B Shah, Kannan Ramchandran & Vijay Kumar, P 2018, 'Information-Theoretically Secure Erasure Codes for Distributed Storage', IEEE Transactions on Information Theory, vol. 64, no. 3, pp. 1621-1646.