International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Hybrid Cloud Storage System with Secure Cryptography using Ecc Algorithm and DataFragmentation Approach

S.Anusuya[1], S.Sushma[2], B.suvetha[3], Prof S.Subashree[4]

[1]*EGS Pillay Engineering College (Autonomous), Nagapattinam, Tamilnadu, India*
E-Mail: anuanu6080@gmail.com
[2]*EGS Pillay Engineering College (Autonomous), Nagapattinam, Tamilnadu, India*
E-Mail: sushmabi1506@gmail.com
[3]*EGS Pillay Engineering College (Autonomous), Nagapattinam, Tamilnadu, India*
E-Mail: suvethabalu04@gmail.com
[4]*Assistant professor, EGS Pillay Engineering College (Autonomous), Nagapattinam, Tamilnadu, India*
E-Mail: subashree@egspec.org

ABSTRACT

Users can access services such as large-scaled data storage and high-performance computing through cloud computing, which connects computer and storage resources managed by different operating systems. Data that is sent to a public cloud must be protected. Clients can use cloud storage services to store data and use high- quality on-demand cloud apps without having to worry about managing their own software, infrastructure, or data. It moves data kept by cloud service providers to cloud storage servers, removing unnecessary burdens for consumers such as physical data possession control. Although the benefits of cloud services are greater, there are new vulnerabilities to data security as a result of physical control of outsourced data. Users are storing sensitive data, and because they no longer have authority over the services or their stored data, strong security techniques are required to prevent unauthorized access to system capabilities and user information. We provide Dispersion and Replication of Data in the Cloud for Optimal Performance and Security, which judiciously divides user files and replicates them at critical cloud sites. The uploaded data was divided into chunks, which were compared to the saved copy and, if a match was found, the unnecessary chunk was replaced with a short reference to the stored chunk. The dispersion of a file into pieces is done according to a set of user requirements, with the individual fragments containing no relevant information. The proxy algorithm ensures that nodes are kept apart. We judicially replicate pieces over the nodes that make the most read/write requests to improve retrieval time even more

Keywords: Cloud computing, Data fragmentation, Replication, Encryption, Cloud security

INTRODUCTION

Cloud computing is unquestionably a potential corporate computing model. It specifies critical infrastructure for a new type of service delivery that has the benefit of lowering costs by sharing computer and storage resources. Cloud computing is currently a massive technology that has surpassed all previous computing technologies in this competitive and demanding information technology market.Cloud computing is rapidly expanding, with several major cloud computing providers such as Amazon, Google, Microsoft, Yahoo, and others offering software-as-a-service (SaaS), platform-as-a-service (PaaS), storage-as-a-service, and infrastructure-as-a-service solutions (IaaS). Furthermore, cloud computing is a fantastic technology since it has the potential to significantly reduce costs through optimization while simultaneously maximizing operational and economic effectiveness. Furthermore, cloud computing has the potential to greatly improve collaboration, speed, and range, enabling a completely global computing paradigm on the internet infrastructure. Furthermore, cloud computing has the benefit of supplying additional scalable, fault-tolerant services. Security risks arise because the data is outsourced to a third-party administrative authority. Data compromise may occur as a result of attacks by other users and nodes within the cloud. In traditional cloud storage, the data owner sends copies of files over the internet to the data server, which records the data to an off- site storage system maintained by a third-party. When the data owner wishes to get information, they use web- based interfaces to connect to the data server. Reliability and security are two issues that exist in the current system. Most systems utilize a variety of approaches to protect data, including: I encryption, which uses a complicated algorithm to encrypt data; (ii) authentication, which requires the creation of a user name and password; and (iii) authentication, which requires the creation of a user name and password. (iii) Authorization, in which the data owner specifies who is permitted to access information saved on the cloud storage system. Even with these safeguards in place, there's still the risk that a hacker will discover an electronic back door and gain access to data. To solve the problem with the current system, the data compression method is used to the cloud storage system, which includes a node allocation process that is meant to secure the data. Because the user is no longer responsible for identifying storage resources, cloud computing handles resource management better. If a user needs additional storage, they can request it from the cloud provider, and when they're done, they can either release the storage by simply ceasing to use it, or relocate the data to a lower-cost long-term storage resource. This further allows the user to effectively use more dynamic resources because they no longer need to concern themselves with storage and cost that accompany new and old resources.

By fragmenting data or files and using numerous nodes to store a single file, the Data Dispersion and Replication methodology prevents crucial data from being leaked. It is advised that you engage a reliable third party to provide cloud security services. The public key infrastructure (PKI) is used in this project to improve the level of confidence in information authenticity, integrity, and confidentiality, as well as communication between the parties involved. The certifying authority are in charge of generating and managing the keys. At the user level, it was suggested that the keys be stored utilizing temper-proof devices made up of smart cards. In cloud environments, we've used general public key cryptography and relied on third parties to provide data security. The authors, on the other hand, have not made advantage of the PKI infrastructure to reduce overheads. The technology and control of public/private keys are the responsibility of the trustworthy third party. A single server or numerous servers might be used as the trusted third party. Combining public key cryptography with (k, n) threshold secret sharing techniques protects symmetric keys. However, such approaches do not protect data files from deterioration and loss as a result of issues originating from virtualization and multi-tenancy.

RELATED WORK

Ren, Xiaoqi, et.al,…[1] Propose a provably optimal algorithm for the case of a data market made up of a single data center and then generalize the structure from the single data center setting in order to develop a near- optimal, polynomial-time algorithm for a geo-distributed data market. The resulting design, Datum, decomposes the joint purchasing and placement problem into two sub problems, one for data purchasing and one for data placement, using a transformation of the underlying bandwidth costs. Given the model of a geo-distributed data cloud described in the previous section, the design task is now to provide an algorithm for computing the optimal data purchasing and data placement/replication decisions, i.e., to solve data cloud cost minimization problem. Unfortunately, this cost minimization problem is an ILP, which are computationally difficult in general. A classic NP-hard ILP is the un-capacitated facility location problem (UFLP). In the un-capacitated facility location problem, there is a set of -clients and J potential facilities. Facility j ∈ J costs fj to open and can serve clients i ∈ I with cost ci,j . The task is to determine the set of facilities that serves the clients with minimal cost. Our first result, stated below, highlights that cost minimization for a geo-distributed data cloud can be reduced to the un- capacitated facility location problem, and vice-versa. Thus, the task of operating a data cloud can then be viewed as a facility location problem, where opening a facility parallels purchasing a specific quality level from a data provider and placing it in a particular data center in the data cloud.

Charapko, et.al,…[2] Propose heuristics to determine the optimal data placement based on the access locality in the workload and load-balancing constraints. Here design and evaluate four data-migration policies following these requirements. Proposed policies work in two phases: first they find an optimal location for some data object given its access history and then adjust this location to adhere to the balancing requirements. Proposed policies work at an arbitrary data granularity, such as individual data items or shards/partitions. They preserve collocation even at the most granular level for related objects having similarities in their access patterns.

Atrey, et.al,…[3] propose a scalable and unified framework for data-intensive service data placement into geographically distributed clouds. The proposed framework introduces a new paradigm for partitioning a set of data-items into geo-distributed clouds using Spectral Clustering on Hypergraphs, and is therefore called SpeCH. Scaling spectral methods to large workloads is challenging, since computing the spectra of the hypergraph laplacian is a computationally intensive task. SpeCH provides two solutions to tackle this problem: (1) an algorithm, called SpectralApprox, that leverages randomized techniques for obtaining low-rank approximations of the hypergraph matrix with bounded guarantees, thereby significantly improving the efficiency of spectral clustering while also providing high quality solutions in practice; (2) an algorithm, called SpectralDist, that exploits the highly parallel nature of the spectral clustering algorithm and uses Apache Spark to speed-up the process while retaining the same quality guarantees as the exact algorithm. Additionally, being distributed in nature, SpectralDist enables SpeCH to perform data placement on workloads that require resources beyond the capacity of a single machine. Since replication may be important in real-world settings for ensuring fault- tolerance and load-balancing, in this section we discuss an extension over SpeCH to allow for the scenarios with replication. Executing the data placement algorithms with different permutations of facilitates data-items to be assigned to different data-centers in each round, thus, ensuring proper replication. Specifically, this procedure allows the same data-item to be stored (in the expected sense) on r data-centers, thereby meeting the desired replication factor. Note that some items might be placed multiple times on the same data-center (of course, in thatcase we only keep a single copy), and hence, each data-item would be replicated at most r times.

Liu, et.al,…[4] Propose Data Bot, a reinforcement learning based adaptive framework, to learn the optimal data placement policies faced with the dynamic network conditions and time varying request patterns. Data Bot utilizes a neural network, trained with a variant of Q -learning, whose input is the real time data flow measurements and whose output is a value function estimating the near-future latency. For rapid decision making, Data Bot is divided into two decoupled production and training components, ensuring that the convergence time of the training will not introduce more overheads to serve the read/write requests. Each server in the system has both the storage and computation functions. For the storage function, data items are distributed among various servers. For the computation function, applications running on multiple servers may require the data movement among servers. All servers are connected through a data center network (DCN). As our objective is to design a generic data placement solution, we do not focus on any specific topology of the DCN. Owing to the fact that our design is only based on the measurement of the end-to-end network performance, it can support any arbitrary DCN topologies, e.g., the tree based Clos and Fat-tree, the recursive DCell and BCube, or the flexible Helios and cThrough architectures. Each data item is assigned with a unique hashtag, i.e., the hash output using the index of the data item as the input. When a data item is written into the system, the metadata server maintains the mapping between the hashtag and its storage server. When an application on an ordinary server needs to retrieve a data item, it first asks the metadata server where the storage server is by using the hashtag. Under this framework, the storage location of a data item is flexible and can be changed, whenever the data item is to be written or updated. By this design, no data movement overhead is introduced even though the proposed framework occasionally changes data storage locations.

Hu, Kekun, et.al,…[5] Propose a two-phase community-aware placement algorithm to place big graphs into the cloud for parallel processing. It can obtain a placement scheme that preserves the community structure well by maximizing the modularity density of the scheme under memory capacity constraints of computational nodes of the cloud in two phases. In the first phase, we design a streaming partitioning heuristic to detect communities based on partial and incomplete graph information. They form an initial placement scheme with relatively high modularity density. To improve it further, in the second phase, we put forward a scale- constrained kernel k-means algorithm. It takes as input the initial placement scheme and iteratively redistributes graph vertices across computational nodes under scale constraints until the modularity density cannot be improved any further. The initial data placement scheme is good but not enough regarding the modularity density.

To improve it further, in the second phase, we present an algorithm named sckernel k-means. It is a generalization of the tradition kernel k-means by adding scale constraints. Given the input π 1,|V| k , it iteratively executes the following two steps, until the modularity density no longer changes: (1) fix the center of each community, and assign each vertex to the community to which its nearest center belongs under the scale constraints; (2) update all community centers. Note that community centers cannot be calculated explicitly because the function φ is unknown. Among these two steps, the first is the key. Essentially, it is a vertex assignment problem. To address it, we propose a greedy-based assignment strategy named GReedy Assignment with Scale Constraints (GRASS). Its basic idea is first to logically assign each vertex to computational nodes of the cloud according to the principle of the shortest distance, ignoring the scale constraints. Then, it adjusts the communities whose scales exceed the corresponding memory capacities of the computational nodes where they are placed.

MODELING AND ANALYSIS

ELEMENTARY

- The cloud is just a mutation form of the Internet. Cloud computing signifies storing andaccessing data and programs over the Internet instead of your computer's hard drive.
- Cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive.
- The cloud is just a metaphor for the Internet. Cloud Computing can be defined as a computer technology that yields the processing power of many inter-networked computers while impersonating the structure that is behind it.

TECHNICALLY

- Cloud computing refers to an efficient method of managing lots of computer servers,data storage and networking.
- The evolution of the term "cloud" can be preferred to the anonymous nature of this technology's framework; the system works for users yet they really have no idea the inherent complexities that the system utilizes.
- Cloud is a new evolution of IT service delivery from a remote location, either over theInternet or an intranet, involving multi-tenant environments enabled by virtualization.

RESEARCH

- Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

EXPERTS

- I have not heard two people say the same thing about cloud. There are multipledefinitions out there of "the cloud".
- {Andy Isherwood, HP's Vice President of European Software Sales} It's stupidity. It'sworse than stupidity: it's a marketing hype campaign**.**

- Everyone who's got an opinion will be telling the world and his dog about theirpredictions for cloud computing.

SERVICE MODELS

To understand broadly Cloud computing has multiple service models like: SaaS, PaaS, NaaS, DbaaS, IaaS, DbaaS and many more. Though every model has its own eminencythe cloud computing has three major types of service models: SaaS, PaaS and IaaS.

SOFTWARE AS A SERVICE

- In simple this is a service which leverages business to roll over the internet. SaaS isalso called as "on-demand software" and is priced on pay-per-use basis.
- SaaS allows a business to reduce IT operational costs by outsourcing hardware and software maintenance and support to the cloud provider.
- SaaS is a rapidly growing market as indicated in recent reports that predict ongoingdouble digit growth.

PLATFORM AS A SERVICE

- PaaS is quiet similar to SaaS rather than SaaS been offered through web the PaaScreates software, delivered over the web.
- PaaS provides a computing platform and solution stack as a service.
- In this model user or consumers creates software using tools or libraries from theproviders.
- Consumer also controls software deployment and configuration settings.
- Main aim of provider is to provide networks, servers, storage and other services.

INFRASTRUCTURE AS A SERVICE

- Infrastructure is the foundation of cloud computing.
- It provides delivery of computing as a shared service reducing the investment cost, operational and maintenance of hardware.
- Infrastructure as a Service (IaaS) is a way of delivering Cloud Computing infrastructure – servers, storage, network and operating systems – as an on-demandservice.
- Rather than purchasing servers, software, datacenter space or network equipment, clients instead buy those resources as a fully outsourced service on demand.

IMPORTANT CHARACTERISTICS

on-demand self-service-

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provide. Possibilities of cloud solutions can be available to the system user in a short period of time, if it is necessary. Let us suppose that our site is in the Cloud and that the traffic,in terms of the number of visitors, is similar every day. Then, let us suppose that one day, for some reason, the Web site traffic rises by 100%. If the is site hosted on our own, private server,there is a strong possibility for it to simply "go down" and stop working because of software and hardware limitations. In such cases, Cloud dynamically allocates necessary resources in order to ensure a smooth operation, and when the flow decreases again, resources are automatically restored to its original condition. The user is free to purchase additional resourcesand opportunities in any quantity and at any time.

WIDE RANGE NETWORK ACCESS

Implies widespread, heterogeneous network accessibility for thin, thick, mobile and other commonly used compute mediums. System capacities are available to customers througha network and can be accessed from different devices such as desktop computers, mobile phones, smartphones and tablet devices.

ALLOCATION OF RESOURCES

Computer resources of providers are grouped in order to serve a large number of simultaneous users. The mechanism of processing power distribution, or the amount of memory, operates in such a way that the system dynamically allocates these parameters according to customer requirements. The users themselves have no control over the physical parameters, i.e. resources location, but at some higher level of the system customatisation, Cloud solutions can choose where their data will be stored and processed (for example, geographical location of data centers).

MEASURED SERVICE

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage,processing, bandwidth and active user accounts). Resource usage can be monitored, controlledand reported, providing transparency for the provider and consumer.

ADVANTAGES

- **Say 'Goodbye' to costly systems:** Cloud hosting enables the businesses to enjoy minimalexpenditure. As everything can be done in the cloud, the local systems of the employees have very less to do with. It saves the dollars that are spent on costly devices.
- **Access from infinite options:**Another advantage of cloud computing is accessing the environment of cloud not only from the system but through other amazing options. Theseoptions are tablets, IPad, netbooks and even mobile phones. It not only increases efficiency but enhances the services provided to the consumers.
- **Software Expense:** Cloud infrastructure eliminates the high software costs of the businesses. The numbers of software are already stored on the cloud servers. It removes the need for buying expensive software and paying for their licensing costs.
- **The cooked food:** The expense of adding new employees is not affected by the applications' setup, installation and arrangement of a new device. Cloud applications areright at the desk of employees that are ready to let them perform all the work. The cloud devices are like cooked food.
- **Lowers traditional servers' cost:** Cloud for business removes the huge costs at the frontfor the servers of the enterprise. The extra costs associated with increasing memory, harddrive space and processing power are all abolished.
- **Data Centralization:** Another key benefit of cloud services is the centralized data. The information for multiple projects and different branch offices are stored in one location that can be accessed from remote places.
- **Data Recovery:** Cloud computing providers enables automatic data backup on the cloudsystem. The recovery of data when a hard drive crash is either not possible or may cost ahuge amount of dollars or wastage of valuable time.
- **Sharing Capabilities:** We talked about documents accessibility, let's hit sharing too. All your precious documents and files can be emailed, and shared whenever required. So, youcan be present wherever you are not!
- **Cloud Security:**Cloud service vendor chooses only the highest secure data centers for your information. Moreover, for sensitive information in the cloud there are proper auditing, passwords, and encryptions.
- **Free Cloud Storage:**Cloud is the best platform to store all your valuable information. The storage is free, limitless and forever secure, unlike your system.
- **Instantly Test:** Various tools employed in cloud computing permits you to test a new product, application, feature, upgrade or load instantly. The infrastructure is quickly available with flexibility and scalability of distributed testing environment.

DISADVANTAGES

- **Net Connection:** For cloud computing, an internet connection is a must to access your precious data.
- **Low Bandwidth:** With a low bandwidth net, the benefits of Cloud computing cannot beutilized. Sometimes even a high bandwidth satellite connection can lead to poor quality performance due to high latency.

- **Affected Quality:** The internet is used for various reasons such as listening to audios, watching videos online, downloading and uploading heavy files, printing from the cloud and the list goes on. The quality of Cloud computing connection can get affected when a lot of people utilize the net at the same time.
- **Security Issues:** Of course, cloud computing keeps your data secure. But for maintaining complete security, an IT consulting firm's assistance and advice is important. Else, the business can become vulnerable to hackers and threats.
- **Non-negotiable Agreements:** Some cloud computing vendors have non-negotiable contracts for the companies. It can be disadvantageous for a lot of businesses.
- **Cost Comparison:** Cloud software may look like an affordable option when compared to an in-house installation of software. But it is important to compare the features of the installed software and the cloud software. As some specific features in the cloud softwarecan be missing that might be essential for your business. Sometimes you are charged extrafor unrequired additional features.
- **No Hard Drive:** As Steve Jobs, the late chairman of Apple had exclaimed "I don't need a hard disk on my computer if I can get to the server faster… carrying around these non- connected computers is byzantine by comparison." But some people who use programs cannot do without an attached hard drive.
- **Lack of full support:** Cloud-based services do not always provide proper support to thecustomers. The vendors are not available on e-mail or phones and want the consumers to depend on FAQ and online community for support. Due to this, complete transparency is never offered.
- **Incompatibility:** Sometimes, there are problems of software incompatibility. As some applications, tools, and software connect particularly to a personal computer.
- **Fewer insights into your network:** It's true cloud computing companies provide you access to data like CPU, RAM, and disk utilization. But just think once how minimal yourinsight becomes into your network. So, if it's a bug in your code, a hardware problem oranything, without recognizing the issue it is impossible to fix it.
- **Minimal flexibility:** The application and services run on a remote server. Due to this, enterprises using cloud computing have minimal control over the functions of the softwareas well as hardware. The applications can never be run locally due to the remote software.

RESULTS AND DISCUSSION

EXISTING SYSTEM

Placing requested data closer to end users helps to minimise the user experienced service delay and the inter-node data read traffic, which encourages significant research regarding data replica placement. Aside from inter-node traffic, the storage locations of data copies might have an impact on the system overhead of accessing related data items. Users can request many data items in a single transaction, which is worth noting. A query may be done by accessing several data blocks in online analytical processing (OLAP) systems, for example. If fewer storage nodes are used to handle such a request, the system overhead could be minimized. The reason for this is that if the read request is sent to a storage node, there will be some overhead, such as the establishing of TCP connections. In a nutshell, data replica placement lowers system overhead by grouping related data items in the same storage location. With the growing amount of data objects, determining the appropriate number and storage location for data copies becomes increasingly important. In geo-distributed cloud storage systems, create scalable and adaptive data replica placement schemes. A data-node community is described as a collection of storage nodes and all data items attached to them, with more internal data access requests than external data access requests. As a result of the smaller community structure, more data requests are serviced locally, resulting in lower system overhead and fewer inter-node traffic. Unlike traditional centralized placement systems, communities can select in a simultaneous and adaptive manner whether each data replica should be placed at the node. Our design's scalability is enabled by this distributed implementation, which has a processing complexity that is proportional to the number of data items.

PROPOSED SYSTEM

In a cloud setting, storing a file in its entirety on a single node creates a single point of failure. If a node is successfully attacked, the data's confidentiality, integrity, or both may be affected. We recommend not storing the full file on a single node in the proposed methodology. The data dispersion technique fragments the file and replicates it on the cloud. The fragments are dispersed in such a way that no single node in a cloud contains more than one fragment, ensuring that even a successful attack on the node exposes on critical data. To boost security, it employs controlled replication, in which each fragment is only copied once in the cloud. A user transfers a data file to the cloud using this method. Upon receiving the file, the cloud manager system (a user-facing server in the cloud that responds to user requests) executes (a) fragmentation, (b) a first cycle of node selection that saves one fragment on each selected node, and (c) a second cycle of node selection for fragment replication. The cloud manager is thought to be a secure entity that maintains track of the fragment placement. To offer secure data storage through encryption, an ECC based encryption technique will be implemented.

PROPOSED ARCHITECTURE

In shows the process of secure file storage in cloud. The owner of the data should upload the files and divide them into fragments. The Blowfish encryption technique was used to encrypt each piece. Following that, encrypted fragments are stored on several nodes. The secret keys are created and stored in the database. Through the request submitting process, a data user can obtain access permission from the data owner. When the data owner agrees to the user's request, the data user is given access to the secret key. Key will have limited access due to time limits. The shared key will be available only during the time period specified by the data owner. Unauthorized data access (whether inadvertent or malicious) by other users and processes must be avoided.. Even after a successful cloud intrusion, the security mechanism must significantly raise an attacker's effort to extract a decent amount of data in such a circumstance. Furthermore, the likely quantity of loss (due to data leakage) must be minimized. The throughput, reliability, and security of a cloud must all be guaranteed. The data retrieval time is a critical aspect in determining the throughput of a cloud that holds data. Data replication solutions are used to solve the challenges of data reliability, data availability, and reaction time in large-scale systems. The following is a description of the ECC-based algorithm:

ECC ENCRYPTION

Input: Parameters from the elliptic curve domain (p, E, P, n), Public Key Q,Raw Text m Output:
Encrypted text (T1, T2) begin

1. Represent the message m as a point M in E(Fp)
2. Select k ∈R[1,n−1].
3. Calculate T1 = kP
4. Calculate T2 = M + kQ.
5. Return (T1, T2) end.

ECC DECRYPTION

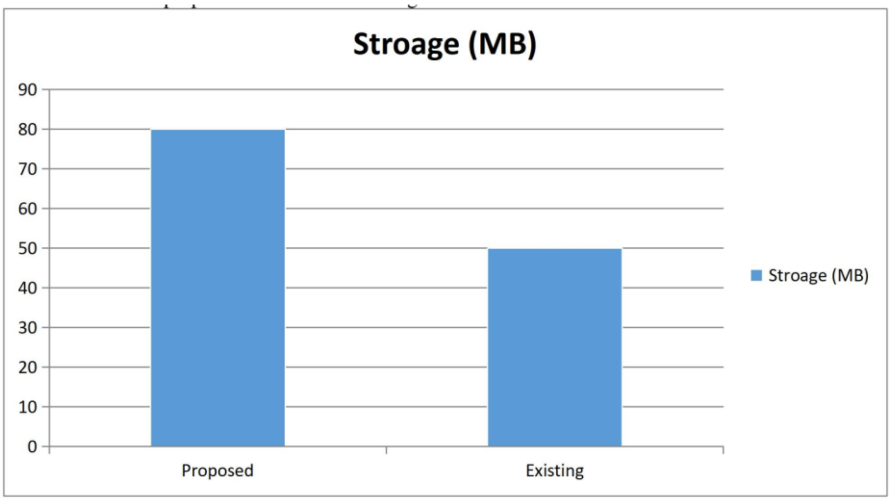
Input: Parameters from the elliptic curve domain (p, E, P, n), Private key d, Encrypted text (T1, T2)Output: Raw Text m begin

1. Calculate M = T2- dT1 and extract m from M.
2. Return (m) end.

In this research, we offer data dispersion and replication in the cloud for optimal performance and security, which addresses the security and execution difficulties in all techniques. We partition a text into portions and repeat the fragmented information among cloud hubs, according to the stated philosophy. Each of the hubs holds only a single piece of a specific information record, ensuring that even if a successful assault occurs, the aggressor is not exposed to any critical data. Furthermore, the proposed approach for information security does not rely on traditional cryptographic algorithms, reducing the need for computationally expensive techniques.

EXPERIMENTAL RESULTS

The proposed algorithm is analyzed in terms of storage preserving and implemented in real time environments. The proposed result is shown in fig 1.



The proposed system preserves the storage up to 80% storage with security.

In shows the two encryption algorithms namely the symmetric and ECC algorithms. They are showing the time consumption of each compared algorithm in seconds when varying the size of data samples.The data sizes are 250MB, 500MB, 750MB and 1000 MB.

CONCLUSION

Secure data storage was implemented using a division and replication scheme in the proposed approach. The user must register in the cloud, and the service provider will send access permissions to each registered user. When a user uploads a file, it is split into small chunks and a secret file key is generated for each upload. When a user wants to download and access a file, they must input the secret file key of their file, after which the split chunks are united and the user can download the file. This ensures security on both the client and the network levels. The future work will concentrate on securing file access using the drops mechanism. This will allow you to focus on your job rather than wasting time and resources downloading, updating, and submitting the content again. Also, concentrate on implementing multiple encryption methods in order to increase performance in terms of encryption and decryption time.

REFERENCES

[1] Ren, Xiaoqi, Palma London, Juba Ziani, and Adam Wierman. "Datum: Managing data purchasing and data placement in a geo-distributed data market." IEEE/ACM Transactions on Networking 26, no. 2 (2018): 893-905.

[2] Charapko, Aleksey, Ailidani Ailijiang, and Murat Demirbas. "Adapting to access locality via live data migration in globally distributed datastores." In 2018 IEEE International Conference on Big Data (Big Data), pp. 3321-3330. IEEE, 2018.

[3] Atrey, Ankita, Gregory Van Seghbroeck, Higinio Mora, Filip De Turck, and Bruno Volckaert. "SpeCH: A scalable framework for data placement of data-intensive services in geo-distributed clouds." Journal of Network and Computer Applications 142 (2019): 1-14.

[4] Liu, Kaiyang, Jingrong Wang, Zhuofan Liao, Boyang Yu, and Jianping Pan. "Learning-based adaptive data placement for low latency in data center networks." In 2018 IEEE 43rd Conference on Local Computer Networks(LCN), pp. 142-149. IEEE, 2018.

[5] Hu, Kekun, and Guosun Zeng. "Placing big graph into cloud for parallel processing with a two-phase community-aware approach." Future Generation Computer Systems 101 (2019): 1187-1200.

[6] D. A. Tran, K. Nguyen, and C. Pham, “S-CLONE: Socially-aware data replication for social networks,” Comput. Netw., vol. 56, no. 7, pp. 2001–2013, 2012.

[7] S. Traverso, K. Huguenin, I. Trestian, V. Erramilli, N. Laoutaris, and K. Papagiannaki, “TailGate: Handling long-tail content with a little help from friends,” in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 151–160.

[8] S. Raindel and Y. Birk, “Replicate and bundle (RnB)–A mechanism for relieving bottlenecks in data centers,” in Proc. IEEE 27th Int. Symp. Parallel Distrib. Process., 2013, pp. 601–610.

[9] R. Nishtala et al., “Scaling memcache at Facebook,” in Proc. USENIX Conf. Netw. Syst. Des. Implementation, 2013, pp. 385–398.

[10] A. Atrey, G. V. Seghbroeck, H. Mora, F. D. Turcka, and B. Volckaert, “SpeCH: A scalable framework for data placement of data-intensive services in geo-distributed clouds,” J. Netw. Comput. Appl., vol. 142, pp. 1–14, 2019.