ISSN: 0369-8963

SECURED REMOTE DATA AUDITING IN DISTRIBUTED CLOUD ENVIRONMENT USING ELGAMAL CRYPTOSYSTEM

MS.K.B.ARUNA[A], DR.C.ARUNACHALAPERUMAL[B], MS.L.SUDHA[C],MS.V.SUREKA[D]

[b]Professor, Department of Electronics & Communication Engineering, S.A.Engineering College,
[a,c,d]Assistant Professor, Department of Computer Science & Engineering, S.A.Engineering College-Chennai,
Corresponding author: Dr.C.ArunachalaPerumal[b]

Abstract-*Managing a large amount of data in cloud is current trend and the challenging task is to secure it against unauthorized manipulations. Cloud computing provides a platform to access data remotely for businesses and individual needs. A remote data integrity check is suggested to ensure the integrity of the data stored in the cloud. The file stored in cloud may contain sensitive information. The private data should not be imparted to unauthorized user when the cloud document is discharged. By scrambling the whole mutual document, Privacy is maintained for confidential information that need to be secured and the personal information can't be viewed or modified by unauthorized person. A remote data integrity checking scheme is proposed that implements information trade with private data that is covered up in the record. A sterilization program is utilized to clean up the data blocks that correspond to the confidential information of document and to change over the signature of the information hinders into legitimate signature for the cleaned record. Subsequently, the plan permits the record put away in the cloud to be shared and utilized by others, provided the secret data is covered up, while the remote data integrity audit can continue to run efficiently.*

Keywords - *Auditing, Data integrity, Legitimate signature, Privacy preserving.*

INTRODUCTION

With the help of cloud computing, users can outsource their private data to a Cloud Service Provider (CSP) and revel in the amazing on-demand services from the cloud. Alternatively, as soon as the proprietors no longer have physical ownership of the outsourced records, the protection of records integrity in cloud computing becomes a vital and difficult task.

A cloud computing affords flexible offerings to the client devices with the aid of which, clients can get entry into the assets those are pooled over the cloud server. Service model used to give an elaborate idea about Storage-as-a-service, by using which clients who have registered to a specific cloud server can region their local data onto the remote storage to be had at the cloud side. A large quantity of data proprietors already started the adoption of Storage-as-a-service to save their neighborhood data onto the cloud server via which they do away with the issue of storage on a neighborhood device and decrease maintenance value. Cloud storage could be very useful provider of a Cloud Computing (CC) by using which facts, proprietors (clients) capable of

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

migrate their nearby data over the faraway cloud server. The faraway servers save their data on cloud servers and customers can get entry into their records from cloud servers whilst implementing the idea of cloud computing. Because of a few security constraints in private data outsourcing, today's concept of information web hosting service,[11] and sensitive information hiding additionally raises new protection demanding situations; those challenges may be handled through third party auditing service to test the data integrity and correctness within the cloud server.

OUTSOURCE AUDIT OBJECTIVES

A. The cloud computing audit objectives:

A cloud audit is a recurring exploration of an organization to estimate and document its cloud vendor's staging. Cloud computing is analogous to the practice of outsourcing of some combination of hardware, software, and data, then to be accessed through Internet connectivity. While a firm intends to adopt cloud computing practice, the primary important task is to pick the proper service provider. Subsequent imperative work is to make a contract with robust Service Level Agreements (SLAs), during which all contractual obligations regarding security, assurance of data systems operations, and data storage should be clearly stated so as to assure the service quality and to assess the danger of cloud computing environment; a firm must seek an audit request through internal or external auditors.

During the audit, the audit team should determine the main target of the auditing project. As indicated earlier, it's going to have value-added or risk–based audit process. Either one should have a selected audit items to be checked. For instance, value-added audit focuses on achieving improved return on cloud computing investment and risk mitigation. On the opposite hand, risk based audit focuses on risk assessment, security, and data safety. The audit team then follows a selected standard or combined standards, frameworks, or guidance to see the compliance criteria. The auditors got to create an audit report during which all checked item should be recorded. These audit items include all SLAs, governance, cost savings, data storage, risk and security issues, and disaster protection on cloud computing's operations within the service provider and industry customer's sites. Basically, there are three sections should be filled, including Relevant Standard's Objective, Audit Procedure, and Findings.

The Relevant Standard's Objective section describes the individual audit standard's objectives that to be audited for cloud computing's compliance. The Audit Procedure section lists all the audit steps and methodologies to be applied into such audit standard's objective. The Findings section reports the auditing outcomes of every audit objective, including positive and negative observations and possible recommendations for improvement. After the audit process, a firm should clearly understand their value of adopting cloud computing environment. Within the meantime, this firm realizes the strength and weakness of moving to cloud computing. Specific risk and security flaws and concerns are often revealed within the audit report.

B. *Principles of Auditing*

The basic principles of auditing is to maintain integrity, audit evidence, skills and competence of management team, accounting system, documentation, planning, confidentiality, work performed by others and internal control, audit reporting to improve its performance of management. Adherence to the subsequent principles are considered to be a prerequisite for ensuring that the conclusions derived from the audit are accurate, objective and sufficient. It also

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

allows auditors working independently from each other to succeed in similar conclusions when auditing in similar circumstances. Auditing procedures include the following,

1. Ethical conduct: Trust, integrity, confidentiality and discretion are essential to auditing;
2. Fair presentation: Audit findings, conclusions and reports reflect truthfully and accurately the audit activities .
3. Professional care: Auditors must exercise care in accordance with the importance of the task they perform;
4. Independence: Auditors must be independent of the activity being audited and be objective;
5. Evidence-based approach: Evidence must be verifiable and be supported samples of the knowledge available.

C. Design goals

To design a verifiable and efficient database auditing scheme which should satisfy the subsequent goals:
• **Privacy:** The data stored in cloud cannot learn any personal secured information that is stored in outsourced database beyond search results, search pattern and access pattern only authorized person can gain access to the data unauthorized access is restricted as specified in policy rules.
• **Correctness:** The cloud cannot forge a fake record which will pass the correctness verification with a clear probability thus originality of the data is preserved and correct data will be maintained.
• **Completeness:** the cloud cannot return a fraction of search results that pass the completeness verification with a clear probability to ensure complete data availability without any damages.
• **Partial Attribute Retrieval:** The client can request the cloud to return only a neighborhood of attributes during a table that satisfy the search condition.

D. Cryptographic algorithms

Many cryptographic algorithms are available to secure the data that can be classified as shown in figure 1, that utilized different cryptographic protocols to make sure data security, database security, and query authentication. Within the following subsections widely used algorithms and techniques utilized in satisfying security requirements are discussed.

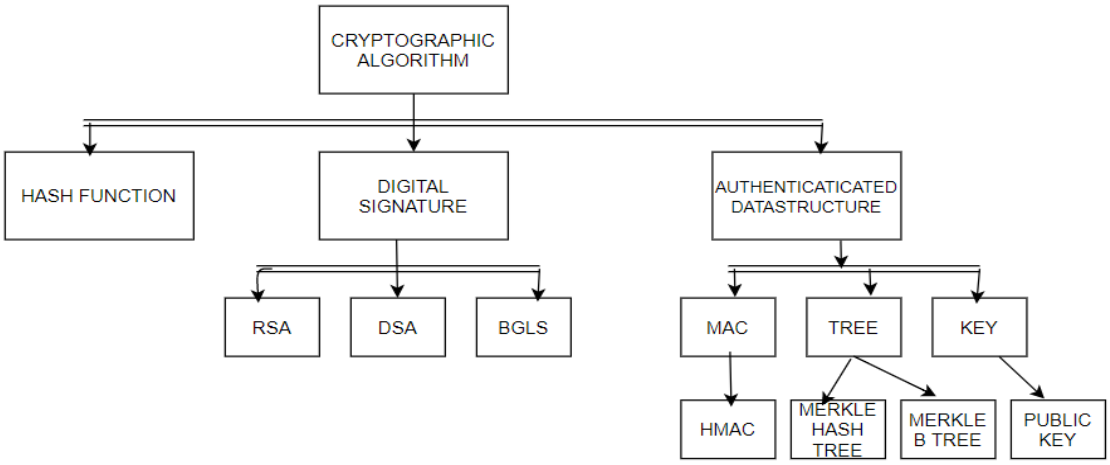

Figure-1 Cryptographic algorithm classification

https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

Related work:

[1] In cloud computing, users can outsource their data to a CSP and luxuriate in the high-quality on-demand services from the cloud. On the opposite hand, once owners do not have physical possession of the outsourced data, the protection of data integrity in cloud computing becomes a critical and challenging task. Idea designed here a is novel verifiable auditing scheme for outsourced database, which may make sure the correctness and completeness of query results albeit the dishonest CSP purposely modifies the results or maliciously returns only a fraction of the results. Leveraging bilinear map and Dynamically Adjustable Capacity cuckoo Filter (DACF) to verify the correctness and completeness of query results. Bilinear map also plays a key role in supporting partial attribute retrieval. Additionally, controllable Paillier encryption and DACF supports data dynamics efficiently. The goal of the proposed DACF is to urge obviate the difficulty of false positive ratio by allocating two new double-sized hash tables whenever an item is to be kicked out. Query integrity refers to the power to verify the correctness and completeness of the query results returned from CSP. Specifically, correctness means the owner must be ready to check the validity of the returned results, i.e., the returned records do exist within the outsourced database and haven't been modified in any way. [13]Completeness means the query results should include all valid records that satisfy the query conditions.

Cloud computing provides flexible services to the client devices by which clients can access the resources those are pooled over the cloud server [2], Service model proposed utilizes Storage-as-a-service, by which clients those are registered to a specific cloud server can place their local data onto the remote storage available on the cloud side. An outsized number of knowledge owners already started the adoption of Storage-as-a-service [15] to store their local data onto the cloud server by which they eliminate the limitation of storage on an area machine and reduce maintenance cost. The prevailing SaaS service providers are Microsoft with Sky Drive, Google Documents and DropBox, etc. They assure data availability and integrity to the user on different systems/locations/networks. Cloud storage is extremely useful service of a CC by which data owners (clients) ready to migrate their local data over the remote cloud server. [16] The remote server (CSP) store their data on cloud servers and users can access their data from cloud servers while implementing the concept of cloud computing. Due to some security constraints in data outsourcing, the newest concept of knowledge hosting service also arises new security challenges; those challenges are often handled by third party auditing service to see the info integrity within the cloud server.

Eight different steps to be followed in a algorithms[3] to check complete data intergity,Setup,DataBlind, AuthGen, DataRecovery, AuthRecovery, AuthVerify, ProofGen and ProofVerify. The protocol's blinding and un-blinding procedures and terms are listed. Therefore, ProofGen and ProofVerify algorithms are an equivalent because the basic protocol used to bind the certificate with the scheme so that security issues and their Impacts are considered and proposed [20] lightweight and privacy preserving cloud auditing scheme by Shen et al. and analyzed its security. By presenting outside adversary attack and cloud server attack, vulnerability of the scheme is concluded. Specifically, both of the surface adversary and therefore the dishonest cloud server can arbitrarily modify data blocks without being detected by the auditor. These attacks showed that despite its positive features, the scheme is insecure and doesn't meet soundness [17] as a basic security requirement of PDP schemes. Next,

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

described a scheme using an improved version of Shen et al.'s protocol which withstands the attacks and also retains other desirable properties of the first scheme. Finally, overhead analysis and experimental results showed the practicality of our improvement.

Dynamic auditing scheme [4] with dispute arbitration describes the thought of index switcher which keeps a mapping between block indices and tag indices. This scheme and show the way to achieve data dynamics support using our index switcher. Finally, we briefly discuss the efficiency of index switcher update caused by dynamic operations.

Modern cloud computing implementation framework establishes more convenient conditions for data sharing in data computing process [5], and provides users with paid network data access protocol. Service providers are required to supply minimal interaction or management services. The overall solution is to transfer the info into the cloud storage system by outsourcing, which has the advantage of realizing the pressure of local processing and storage and realizing lower data maintenance cost.In order to enhance the computational efficiency, an algebraic signature Divide-and-Conquer Table (DCT) cloud storage remote auditing method for the large data is proposed. The contributions: (1) A high-efficiency cloud computing data storage remote auditing scheme supported the algebraic signature is developed. This solution enables a discount within the computing and communication costs of the auditor and server. (2) A replacement arrangement, DCT, is meant to frequent updates of large-scale data, achieving minimal computing costs of auditors and servers. (3) The remote data auditing method is implemented within the real environment. It's proved that it provides better data security and performance than the foremost advanced data auditing methods. Subsequent step is to theoretically and experimentally analyze the safety of the algorithm,

Proposed a mechanism with the CSP may plan to pass the verification using the proof generated from the previous ones or other former information [6],

1. Public auditing: anyone (not only the users) is allowed to possess the potential to verify the correctness and integrity of the users′ data stored within the cloud.
2. Storage correctness: the CSP, which doesn't correctly store users′ data as needed, cannot pass the verification
3. Block less verification: no data block must be retrieved by the TPA during the verification process.
4. Dynamic data auditing: dynamic data operations should be supported while the efficient public auditing is achieved.
5. Privacy preserving: the TPA cannot derive any actual content of users′ data from the received auditing information.
6. Batch auditing: the TPA can handle multiple auditing tasks from various users during a fast and cost-efficient manner.

Data-integrity verification [7](by a 3rd party auditor) for the client's data residing on a cloud storage server (CSS). Also optimize the existing third-party auditing protocol and make it immune to replace, replay and forge attacks launched by malicious insiders at cloud storage server. They propose a protocol to perform efficient block-level and fine-grained dynamic-data update operations on data stored on cloud employing a modified Chameleon Authentication Tree. Optimized the general public auditing of client's data residing on cloud by shifting the HLA's from CSS site to TPA's site. Introduced a protocol for performing dynamic data

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

update employing a modified Chameleon Authentication Tree (mCAT). However proved the safety of optimized auditing protocol by showing that it's immune to replay, replace and forge attacks. we've shown the computation cost(time) for generation of mCAT at the client's side and overall computation time taken for block authentication and update of authentication path during dynamic data update phase. The Computation time for each auditing task is specified in auditing scheme for shared data to achieve the shared data integrity along with user revocation in the cloud.

A mechanism to design a foreign data audit mechanism [8] the subsequent important criteria must be taken into account: (1) **Efficiency:** audit the info with the minimum computational cost over the server and particular client. The auditing service is additionally reasonable for the communication overhead between client and server, (2) **Public verifiability**: delegate the audit task to a trusted third party auditor instead of a client so as to scale back the computation cost over the client,(3) **Frequency:** number of times that user is in a position to verify the integrity of outsourced data by generating a challenge message,(4) **Probability of detection**: probability by which a protocol detects data corruption,(5) **Recovery:** ability to recover data just in case of knowledge corruption (6) **Dynamic update:** enabling the cloud user to update the outsourced data by using insert, delete, modify, and append operation without requiring to download the entire data. The following security patterns are practiced to audit the outsourced data within the current data storage security schemes:

(1) **Homomorphic encryption**: encoding may be a crucial method to store and access data securely within the cloud. However, the most issue is the way to perform computation on encrypted data received by the cloud server without having to decrypt it and to get an equivalent result as working on the first data. Rivest was the primary to realize this goal by proposing a homomorphic encryption. In single cloud server, the homomorphic encryption mechanisms are categorized into the subsequent two types:

(i) **Homomorphic verification tag (HVT**) allows the client to mix the computed tags for multiple blocks of every file into one value

(ii) **Homomorphic Linear Authentication (HLA**) utilizes a linear combination of the individual data block to get one value. Since HLA uses a comparatively small-sized BLS legitimate signature, it incurs less computation overhead than HVT

(2) **Pairing-based cryptography:** the most idea behind this method is to construct a mapping between the weather of two cryptographic groups for building a cryptosystem on the idea of the reduction of 1 problem in one group. It is often included Cryptographic Bilinear Pairings and Diffie–Hellman assumption

(3) **Symmetric Key Cryptography:** during this sort of cryptography, a shared secret key's won't to encrypt the plaintext and decrypt the cipher text to protect the data against the unauthorized user while tampering the outsource resource without damaging the original data asset.

(4) **Polynomial commitment scheme:** Helps a committee to plan to a polynomial with a brief string which will be utilized by a client to verify claimed evaluations of the committed polynomial.

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

(5) **Computational Diffie–Hellman (CDH)** may be a valuable assumption for cryptographic purposes and relates to the problem of computing the discrete logarithm problem within a cyclic group .CDH cares with the mathematical operations that are completed quickly, but difficult to reverse.

The two categories of verifier's role that are: [9] first one is private auditing, during which only customer or information proprietor is allowed to verify the honesty of the hoarded information. No other person has the authority to question the server regarding the information. But it tends to extend verification overhead of the user. Second is public auditability, which allows anyone, not just the customer, to challenge the server and performs information verification with the assistance of TPA [19]. Distributed repository verifying protocols with built-in key disclosure resilience is meant that reduces the damage of the customer's secret key disclosure. The algorithms support forward security and property of block less verifiability which incurs high overhead to accomplish extra key disclosure resilience. Public verification for reconstructing code based distributed repository achieves protection of stored data against exploitation and adds fault tolerance to distributed repository. The mechanism relieves clients from online burden. There's a requirement to style protocols that support optimal error correcting codes.

Intrusion resilient public cloud auditing scheme [10], in this scheme while auditing the data, authenticators will update periodically to stop the malicious cloud from tampering the available files using the exposed key. This scheme is secured unless the client and TPA (Third Party Auditor) are compromised during auditing phase, data can be accessed dynamically among the multiple group user in same time. Security goals consists of data proof consistently:

(1) Intrusion-resilient: The scheme of cloud storage auditing is intrusion-resilient, if an adversary A can make the challenger C accept its proof with a probability which is non-negligible, there's an efficient algorithm to get the challenged file blocks with a non-negligible probability to verify the integrity of the original valid data of data owner.
(2) Privacy-preserving: The scheme of cloud storage auditing using privacy-preserving technique, is a curious when TPA cannot recover the client's files supported proof P returned by the cloud server hence additional care need to be provided for outsourced data.
(3) Detectability: The scheme of intrusion-resilient public cloud auditing is (q, p) detectable, for corrupted blocks which have a fraction q, the probability that these corrupted blocks are going to be detected isn't but p. The integrity verification guarantees the integrity of fingerprint [22], in online transmission and gives the digital signature step an accurate fingerprint. Signature verification checks the fingerprint integrity by retrieving the developer public key from the developer public keys database. CA signs the plaintext of the fingerprint to generate the digital signature using the private key contained in CA's key database once the fingerprint's integrity has been validated. The received fingerprint is validated to generate the encrypted fingerprint, which is subsequently provided to its related developer during the fingerprint authentication phase.

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

Proposed Scheme:
Lightweight secluded data auditing scheme (LSDA) is used to ensure data integrity of outsourced data with the minimum communication time and computation cost.

Forge attack: The CSP may forge the info blocks and/or their tags to deceive the verifier.
Replacing attack: The CSP might want to pass the verification by replacing a required block and its tag, which are corrupted, with another block and its corresponding tag.
Reply attack: The CSP may plan to pass the verification using the proof generated from the previous ones or other former information.

Cloud users not physically possess their data, so the way to make sure the integrity of their outsourced data becomes a challenging task. [14]Challenge schemes like "provable data possession" and "proofs of retrievability" are designed to audit static archive data and don't provide knowledge about dynamics support data manipulation operations. Moreover, threat models in these schemes usually assume an honest data owner and specialize in detecting a dishonest cloud service provider despite the very fact that clients can also misbehave. Firstly, despite the powerful machines and powerful security mechanisms provided by CSP, remote data still face network attacks, hardware failures and administrative errors. Secondly, CSP may reclaim storage of rarely or never accessed data, or maybe hide data loss accidents for reputation reasons. As users not physically possess their data and consequently loose direct control over the information, direct employment of traditional cryptographic primitives like hash or encryption to make sure remote data's integrity may cause many security loopholes. Especially, downloading all the info to see its integrity isn't viable; thanks to the expensive communication overhead, especially for large-size data files. During this sense, Message Authentication Code (MAC) or signature based mechanisms, while widely utilized in secure storage systems, aren't suitable for integrity check of outsourced data, because they will only verify the integrity of retrieved data and don't work for rarely accessed data (e.g., archive data). So, the way to make sure the correctness of outsourced data without possessing the first data becomes a challenging task in cloud computing, which, if not effectively handled, will impede the wide deployment of cloud services. Data auditing schemes can enable cloud users to see the integrity of their remotely stored data without down- which is termed as block less verification. With auditing schemes, users can periodically interact with the CSP through auditing protocols to see the correctness of their outsourced data by verifying the integrity proof computed by the CSP, which offers stronger confidence in data security because user's own conclusion that data is unbroken is far more convincing than that from service providers

Data Center: The data owner should be a business user who understands the financial implications of a security incident that results in a loss of availability, confidentiality, or integrity. Data owners are in charge of determining who has access to particular system functions and datasets, as well as what they may do with the information.

Auditor: A cloud auditor is a third party who can conduct an unbiased review of cloud service controls with the goal of expressing an opinion. Audits are conducted to ensure that standards are being followed by reviewing objective evidence. A cloud auditor can assess a cloud provider's services in terms of security measures, privacy implications, and performance.

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

CSP: CSP is a company that manages all aspects of networking, software, servers, infrastructure, and other services. They also hire and manage employees, as well as providing protection for the services they supply.

End user: End user is authorized to do one or more of the following things with the data: read only, update, create, and remove Individual roles can be defined using role-based access controls (RBAC). The data owner must be satisfied that the contractual relationship with the CSP is backed up by suitable controls that show that what was expected to happen actually happened and that the outcome met the contract's requirements. During the evaluation of proposed CSP's history of data breaches and its accompanying reports, the data owner should perform due diligence.

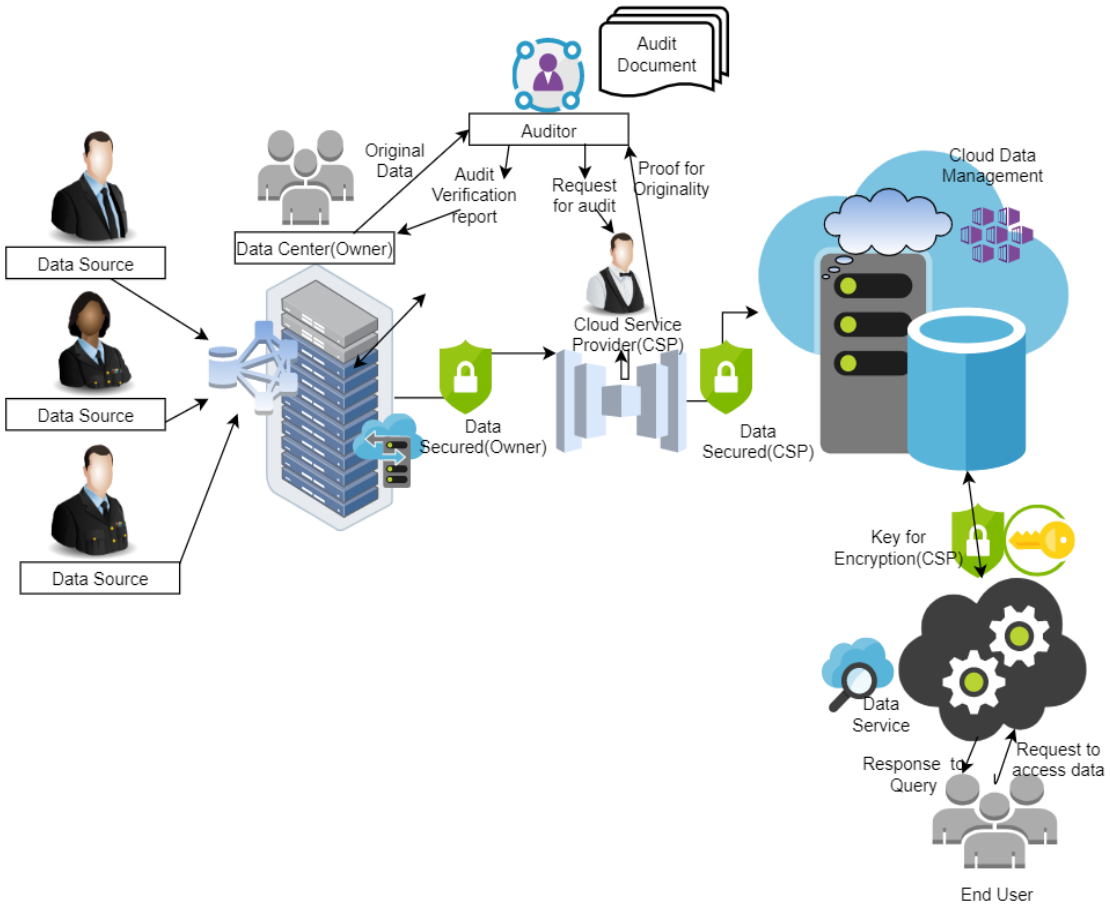

Figure-2 Architecture to Secure Remote Access of Resources

Figure 2, Describes the various phases to be followed to ensure data integrity in outsourced environment,

Phase 1- Data's from various data sources like workstation, service level agreements, policies, protocols, European Union Open Data Portal. Business Analytics Work Center, File Data Source or Machine Data Source, a technical database, predictive data which includes historical, current, future trend of data is collected and maintained by data owner as original source of data.

Phase 2- Data owner handover the secured data to the trusted CSP based on service level agreement.

Phase 3- CSP shares the data with end-users on demand basis.

https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

Phase 4- Dataowner shares his original data to Auditor and assign the task of data integrity challenge provided by CSP on contract basis, Auditor tracks CSP for challenge, CSP as a result submit the proof for verification. Now it's the responsibility of the auditor to prepare the template and analyze the detailed report of proof of the data verified. Followed by, report will be given as response to the request assigned by the DO. Streamline resource management, increases operational efficiencies, optimizes product development, drives new revenue and growth prospects, and the final audit report should be supportive for decision making process for the management about their business

Setup & policy maintenance in phase1:

Data lakes retain a large volume of raw data in its original formats. When a data lake is searched, a subset of data is picked based on search criteria and presented for analysis; when a data lake is queried, a subset of data is selected based on search criteria and presented for analysis. The Snowflake platform combines the advantages of data lakes with those of data warehousing and cloud storage. Organization will benefit from best-in-class performance, relational querying, security, and governance using Snowflake is core data store, like data's stored in Amazon S3 or Azure. To speed up data converters and analytics, use Data Lake cloud storage and Snowflake. The goal of a data lake is to make organizational data from various sources accessible to various end-users such as business analysts, data engineers, data scientists, product managers, executives, and others so that they can leverage insights in a cost-effective manner for improved business performance. Many types of sophisticated analytics can now only be performed on data lakes. A data lake allows data to be accessed using a standard-based SQL implementation that does not require any proprietary modifications. It uses standards like ODBC and JDBC to allow external tools to access the data. Additionally, conventional programming languages such as R, Python, and Scala, as well as standard libraries for numerical computation and machine learning such as TensorFlow, Apache Spark, MLib, MXNet, Keras, and SciKit Learn, provide programmatic access to data in a data lake.

Setup & agreement maintenance in phase 2:

New cloud data governance systems and techniques must be adopted by data architects and data engineers to enable efficient data stewardship, assist innovation by shortening time-to-data, and automate compliance with data privacy requirements. Modern tools that empower data engineers and compliance teams to automate data governance, data access rules, and privacy protection – all from a single software interface – can help data teams negotiate the complexities of cloud data governance. Even if sensitive data discovery is automated, data governance teams must be able to validate that it has been correctly recognized, categorized, and tagged. Data architects and engineers should create workflows for checking, assessing, and approving the findings of automated discovery and tagging to meet these needs. Unauthorized data consumers can be denied access to specific rows, columns, or cells in a table using fine-grained data access rules. Fine-grained data access controls enable enterprises to comply with data requirements while also protecting sensitive data that is stored in a table alongside other commonly used data or that must be accessed for a specific purpose. Immuta can be integrated with Snowflake to enable more granularity, scalability, and automation than the native controls in Snowflake. Data teams may use Immuta's close interaction with Snowflake to apply Snowflake's basic row access and column masking policies while also boosting security using Immuta's highly scalable

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

Attribute-Based Access Controls (ABAC) and other additional capabilities. Immuta provides sophisticated security, access control, auditing, and privacy to Snowflake customer's management. using Snowflake as your primary BI platform, you'll need a solution that allows you to mask data across Snowflake and any other platform in your data stack. Immuta meets this demand with centralized, universal data access control, sensitive data detection and classification, and consistent data masking, as described in this article. You can construct a global masking policy to apply dynamic data masking across all fields in Snowflake and any other platform using Immuta's policy-as-code capabilities. Hashing, regular expression, rounding, conditional masking, replacing with null or constant, reversibility, format preserving masking, and k-anonymization, as well as external masking, are all examples of this. The policy applies to "everyone except" individuals with the attribute "Department" set to "Human Resources," which is derived from an attribute set in external system. This dynamic technique, also known as attribute-based access management, can reduce roles by a factor of 100, making data more manageable and lowering risk for data engineers and architects. AWS, Microsoft Azure, Google, and other cloud service providers have made it easier for their users to meet security guidelines, criteria, and certifications by making it simple to set controls that auditors will be searching for. Furthermore, many companies include a wealth of information in their white papers so that users may determine whether their product meets the security criteria.

Setup & protocol checkup in phase 3:

Cloud data governance policies must be audited on a regular basis to determine the effectiveness of existing procedures, detect any security risks or threats, and ensure continuous regulatory compliance. Data architects and engineers should provide capabilities to monitor and log data usage to support the data audit trail. Data-rich audit logs that include all data sources, who subscribes to each one, when they were accessed, what data was accessed, and all queries executed enable data teams to share data usage details with compliance and legal teams, and are critical for proving compliance and troubleshooting issues when needed. Immuta also offers row-level filtering and dynamic privacy-enhancing technologies (PETs), such as differential privacy and randomized response, in addition to column restrictions. The native query that was executed, such as: handler – This is the type of native integration. startTime – the time the query began in UTC, endTime – the time the query ended in UTC, duration – the time the query took in milliseconds, nativeObject – The object that was requested in its entirety. nativeObjectType – The object that was queried's type (e.g. view or table), host – This is the host to which the native integration is linked. Database -This is the database where the native object is stored.

A Bloom filter is a space-efficient probabilistic data structure for determining whether or not an element belongs to a set. False positive matches are conceivable, but false negative matches are not a query returns "maybe in set" or "certainly not in set." Elements can be added to the set but not removed (though the counting Bloom filter variation can help with this); the more items added, the higher the chance of false positives. The technique was proposed for applications where the amount of source data would necessitate an impractically huge amount of memory if "traditional" error-free hashing algorithms were used. When there is enough core memory, an error-free hash can be used to remove all unnecessary disc accesses; when there isn't enough core memory; Bloom's technique uses a smaller hash area but still eliminates most unnecessary accesses. Multiple Bloom filter layers make up a layered Bloom filter. By verifying how many layers include the item, layered Bloom filters allow you to keep track of how many times an item was added to the Bloom filter. A check operation on a layered Bloom filter will

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

usually return the item's deepest layer number. A skip list is a probabilistic data structure that enables both search and insertion complexity within an ordered succession of elements. As a result, it can have the greatest features of a sorted array (for searching) while keeping a linked list-like structure that allows insertion, which is impossible in an array. Maintaining a linked hierarchy of subsequences allows for faster search, with each consecutive subsequence skipping fewer entries than the previous one. The search begins in the sparsest subsequence and continues until two successive elements, one smaller and one larger than or equal to the element searched for the linked hierarchy, are located.

Setup & report generation in phases 4:

Immuta tracks and logs all actions on your data platform, making it easy to demonstrate compliant data usage and investigate incidents. Furthermore, because access control policies can be represented as code, Software Development Life Cycle principles can be used to policy change management, automating DataOps continuous delivery procedures. Immuta policies can be updated to implement controls across many Snowflake accounts using editable templates. Auditor Roles: Activity Monitoring Examine documentation that identifies system flaws, and examine system configurations to see if notifications are sent out when flaws or failures are discovered. Observe whether the office requires a badge to enter, examine proof that individuals with administrator level access are authorized, and examine the password policy used to access the network. Operation of Systems Examine the monitoring tools that are used to monitor traffic and alert on suspicious activity. Examine evidence that the tools are able to provide alerts as required. Examine evidence that notifications are followed up on and resolved as needed. Change Management-Inspect evidence to ensure that changes have been defined and documented, that they have been approved for development, that they have been tested, and that they have been approved for implementation. The records or documentation of procedures done by auditors, the audit evidence acquired, and the conclusions reached based on the evidence obtained are referred to as audit documentation. The term "audit working" is sometimes used to refer to audit paperwork. Examine evidence that stated flaws have been addressed.

Security Analysis at each phase:

Phase1: A data lake is a centralized repository that allows structured and unstructured data to be stored at one place, with any scale. It is integrated with features like dashboards and visualizations to make better decisions; it also helps to run several sorts of analytic from big data processing, real-time analytics, and machine learning without need to organize the data. While any sample assessment is tailored organization's size, archives, also known as records or record office, is a repository for an organized body of records created or received by a public, semipublic, institutional, or business entity in the core of its business and preserved by it or its successors. The term archives, which also refers to the collection of records as a whole.

Phase 2:

To secure the data which is afford on demand basis by CSP ,following steps are used at first cryptosystem concept to secure the data is applied a as follows, Microsoft public cloud subscriptions, such as Office365, Enterprise Mobility Suite (EMS), Azure, and Dynamics CRM Online, are managed through CSP as it is licensing platform. Next enrolling the company's

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

business verification validates legitimate business entity with the specified address. The CSP program allows clients' operations that benefits: Customer utilization of resource more in-depth:. So that CSP can gain more Profits on support and billing services, via directly or through a third-party source, that creates additional revenue streams. Business-specific solutions combined with Microsoft products, to meet client demand for managed services.

ElGamal Cryptosystem

Like RSA, the public-key cryptosystems classified and used based on different versions of the Discrete Logarithm Problem.

ElGamal cryptosystem, is also called Elliptic Curve Variant, its logic depends on the Discrete Logarithm Problem. It procures the robustness from the assumption that the discrete logarithms cannot be found in practical time frame for a given number, while the inverse operation of the power can be computed efficiently.

The basic variant of ElGamal works with numbers modulo p. In the case of elliptic curve variants, it is based on quite different number systems.

Generation of ElGamal Key Pair:

Each user of ElGamal cryptosystem generates the key pair through as follows −

- Choosing a large prime p. Generally, a prime number of 1024 to 2048 bits length is chosen.
- Choosing a generator element g.
    - Choose a integer value between{ 1,.., p − 1 }
    - It is a generator of the multiplicative group of integers modulo p. This means for every integer m co-prime to p, there is an integer k such that gk=a mod n.
      For example, 3 is generator of group 5 (Z5 = {1, 2, 3, 4}).
- Choosing the private key. The private key x is any number bigger than 1 and smaller than p−1.
- Computing part of the public key. The value y is computed from the parameters p, g and the private key x as follows −y = 65 mod 17 = 7

Encryption and Decryption

The generation of an ElGamal key pair is comparatively simpler than the equivalent process for RSA. But the encryption and decryption are slightly more complex than RSA.

ElGamal Encryption

Suppose sender wishes to send a plaintext to someone whose ElGamal public key is (p, g, y), then −

- Sender represents the plaintext as a series of numbers modulo p.
- To encrypt the first plaintext P, this is represented as a number modulo p. The encryption process to obtain the ciphertext C is as follows −
    - Randomly generate a number k;
    - Compute two values C1 and C2, where −C1 = gk mod p
    - C2 = $(P*y^k)$ mod p

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

ElGamal Decryption

- To decrypt the ciphertext (C1, C2) using private key x, the following two steps are taken −
    - Compute the modular inverse of (C1) x modulo p, which is (C1) -x, generally referred to as decryption factor.
    - Obtain the plaintext by using the following formula −

C2 × (C1)-x mod p = Plaintext

Polymorphic viruses:

Polymorphic viruses hide their presence through a series of encryption and decryption cycles. A decryption program first decrypts the encrypted virus and its accompanying mutation engine. The virus then infects a section of code. The virus encrypts the mutation engine and a copy of the virus with an algorithm corresponding to the new decryption procedure, and the mutation engine produces a new decryption routine. The mutation engine and virus's encrypted package is attached to new code, and the process is repeated. Because of the numerous modifications to their source code, such viruses are difficult to detect but have a high amount of entropy. This characteristic can be used to detect them by anti-virus software or free program like Process Hacker. This type of attacks can be prevented and secure the access with restricted modification of data.

Phase 3: Ratings are based on variables such as vulnerabilities, compromised systems, adherence to industry best practices, and compliance with cyber security guidelines. The results are presented as a simple numerical score, with a higher score indicating better overall security performance. If a vendor's rating is low, you may decide not to sign a cloud services deal with them. You can also work with them to enhance their rating if you deem them business-critical. You need a mechanism to independently assess risk based on data-driven insights — from onboarding through the life of the relationship. Using a service like security ratings, you can easily automate this procedure.

Data integrity includes three aspects: – Correctness: the query issuer can check that the returned results are correct. - Completeness: the final product is comprehensive, with no missing answers. - Reliability: the results are based on the most recent data version, an act or process in which an unauthorized person or resource tries to gain access to another person's data without their knowledge or consent. It's just a security breach in which data is accessed without authorization. It's possible that various people have different ideas about how to get access to such data without permission. Digital signature verification process as follows as shown in the figure3,

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

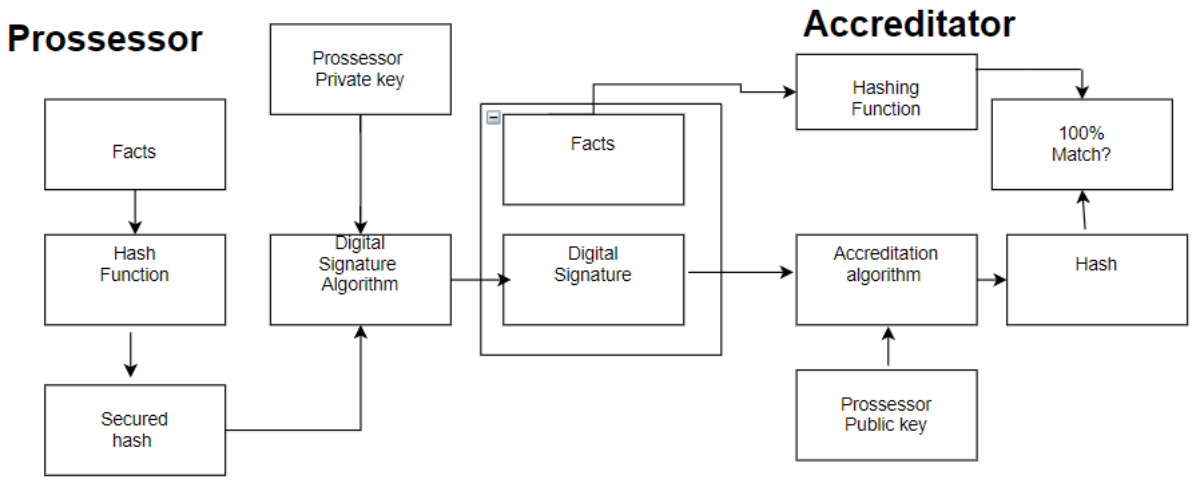

Figure 3: Digital Signature verification Process.

[1] Verifiable and efficient auditing scheme for privacy preserving of outsourced database without a third-party auditor (TPA) is the proposed scheme that simultaneously authenticates the correctness and completeness of query results, while also preventing them from being tampered with dishonest CSP from returning fake records or only a fraction of the query results to the users. It also allows variable data dynamics and partial attribute retrieval, which reduces transmission costs significantly. Bilinear maps are used to achieve correctness verification and facilitate partial attribute retrieval in particular. A DACF without false positive ratio is proposed and built by different attribute columns to realize completeness verification; together with DACF, a controllable Parlier encryption is also proposed to support data dynamics including inserting, deleting and updating data. Finally, proposed scheme helps to handle the remote data to protect it against unauthorized access, its communication cost, its integrity verification cost, and dynamic data operation cost, reduced cost in deleting a record from database. [2] Optimized secure dynamic auditing protocol is secure and efficient against various conspiracy attacks.

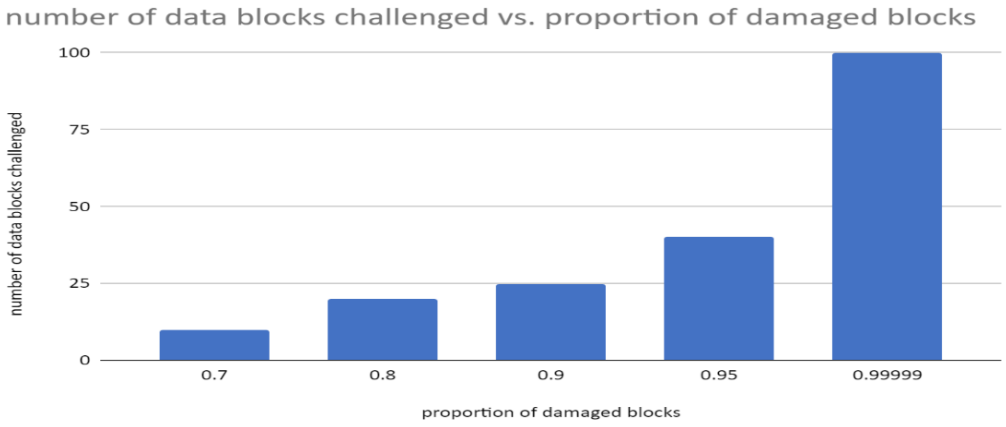

Figure 4: Data block challenge

https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

Phase 4: CCM v4 (Cloud Controls Matrix) is one of the cloud security metrics that intended to enhance internal CSP governance, risk, and compliance (GRC) efforts and serve as a useful benchmark for service-level agreement disclosure. Analyze. Examine how effectively the vendor's practices adhere to CSA and ISACA guidelines. Compile your findings. Work papers are used to generate a final report and recommendations by combining analysis with evidence from documentation and interviews. Complete the final report. It should be presented to the organization's management during a formal audit briefing. Make a move. Management assigns a team to respond to the audit report and establishes dates for replies to the suggested actions. The evidence obtained by the auditor, the techniques used for testing, and the results of testing should all be documented appropriately and clearly in the audit working papers as shown in figure4 data block challenge graph which considers the parameters like number of data blocks challenged and proportion of damaged blocks timely representation of data will be presented. This is to ensure that the quality reviewer can easily complete the task and that the applicable standards are being followed. After negotiating with the customer, the auditor will conduct the final audit, and they will agree on a time frame for the audit.

Performance efficiency:

The main focus will be on pseudonym-based auditing and protecting auditing storage against various assaults. The process activities based on the formation of traces, store the trace beginning data in the header so that running processes are not slowed down. Simulative analysis can be used to demonstrate dense network testing.

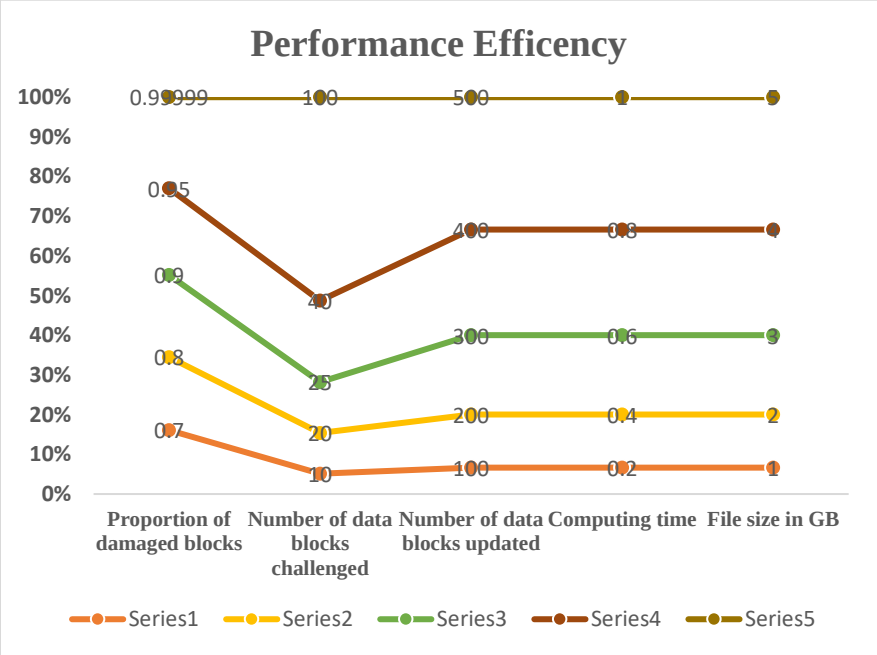

Figure 5: Performance efficiency.

[3] Lightweight and privacy preserving cloud auditing scheme analyzed its security scheme's vulnerability by providing an outside adversary assault and a cloud server attack. Both the external adversary and the dishonest cloud server can alter data blocks arbitrarily without

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

being discovered by the auditor. These assaults revealed that, despite its benefits, the plan is insecure and fails to meet the soundness criteria of PDP schemes. [4] Provides a public-verifiable integrity auditing scheme with efficient data dynamics and fair dispute resolution. The target is to Identify the variation in block indices and tag indices, that introduced in index switcher to maintain originality of block-tag index mapping to reduce tag re-computation which is caused by block modification/new entry operations, which incurs reduced overhead, as shown in our performance evaluation. Meanwhile, because both clients and the CSP may misbehave during auditing and data manipulation hence a threat prediction model is introduced that enables fair arbitration for resolving conflicts between clients and the CSP. [5] In order to improve the computational efficiency, an algebraic signature DCT cloud storage remote auditing method for the big data is proposed. The contributions: (1) a high-efficiency cloud computing data storage remote auditing scheme based on the algebraic signature is developed. This solution enables a reduction in the computing and communication costs of the auditor and server. (2) A new data structure, divide and-conquer table (DCT), is designed to frequent updates of large-scale data, achieving minimal computing costs of auditors and servers. (3) The remote data auditing method with data security is applied to the real time data set.

[6] Minimize the handling cost during the verification process of original data tampering the security of our scheme of auditing process with experiments and comparisons with the existing ones is detailed analyzed The results effectively achieve secure auditing in clouds, and induce significantly fewer costs of storage, communication, and computation than the previous schemes also no single method can achieve perfect audits for all types of cloud data, just as no standard has a universal verification and validation thus different audit strategies for various types of cloud data, [7]For provocation generation + retort generation + probity verification in turn number of security scheme reviewing demonstrated the security that is detailed and streamlined inspecting convention by showing that it is unable to replay for attacks, alter and produce false results also computing time its complexity mCAT at the end user side and in general performance efficiency which is exact time taken for block validation and update of verification way during dynamic information update stage how our evaluating plan works and supports the process. [8],[12]data quality management, Migrating computational functions along with data into the cloud and using challenge–response approach to check calculation and information trustworthiness can be a potential method to resolve this issue. This procedure, which is valuable for asset limitation gadgets to decrease the calculation cost actually needs specific level of consideration. A few open difficulties especially, lightweight information reviewing, dynamic information update, information access control, and computational trustworthiness.

[9] Proficient examining of cloud consistency is accomplished, and the clients can definitively pick the cloud specialist co-op. Proficient calculations should be planned that work on the capacity and calculation overhead of the current algorithms. The calculations support forward security and property of block less evidence which causes high overhead to achieve additional key divulgence flexibility. Public check for recreating code based dispersed vault accomplishes assurance of put away information against abuse and adds adaptation to internal failure to circulated storehouse. The system eases customers from online weight. There is a need to plan conventions that help ideal mistake amending codes. [10] An intrusion-resilient public cloud auditing scheme with authenticator update, which adequately diminishes the damage brought about by key-openness. The plan refreshes the authenticators [18] in each timeframe and takes care of the issue that the cloud worker messes with documents which are transferred in the

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

key-openness time frame. On the off chance that the customer and TPA are compromised in various time spans, the customer's reviewing secret key is as yet secure. Furthermore, the plan ensures the customer's record security to forestall inquisitive TPA from recuperating the customer's documents. With the security examination and the consequences of the test, it shows that the security and execution of our plan are acceptable. Although this paper has proposed an interruption versatile public cloud reviewing plan, there are still a few issues worth considering later on. For instance, how to ensure the security when the customer and TPA are compromised simultaneously can be worked on later on. What's more, it is beneficial to stretch out the plan to the IoT climate by consolidating edge processing.

Conclusion:

In this paper various Auditing techniques discussed that gives an clear idea about Auditing concept which is applied to cloud computing, mobile cloud computing which means service level agreement, Several open challenges particularly, lightweight data auditing, dynamic data update, data access control, and computational integrity were presented that can be done between the server and client in order that client can get help from TPA for effective verification and security of the data stored on the remote cloud server. Clearly the fundamental properties for implementing TPA must state the dynamic auditing protocol the main target is to save auditing computation against different attacks. Within the auditing process, the communication between the auditor and remote server is completed through two-way communication where auditor sends challenge to a foreign server and in response server sends the proof after accessing the challenge by this auditor ensures data is correctly stored at cloud side. Further auditing data are going to be updated consistent with the specified storage solutions. After the confirmation of auditing process, TPA will send the result to the info owner. If the result sent by the TPA is true, then data owner is convinced that data remotely stored over cloud server is correctly stored. Then data owner may prefer to delete the local copy of knowledge. In turn general public key, private key and a certificate to take care of privacy, users define their candid and certificate validate that user may be a regular user. During the auditing process Constant bandwidth cost, Protecting data privacy, Batch auditing Data owner, data dynamic support, Low computation complexity, storage costs and communication overhead need to be ensured. How to guarantee the safety when the client and TPA are compromised simultaneously are often improved within the future.

References:

[1].Tao Xiang a, *, Xiaoguo Li a , Fei Chen b , Yuanyuan Yang c , Shengyu Zhang, J. “Achieving verifiable, dynamic and efficient auditing for outsourced database in cloud” Parallel Distrib. Comput. 112 (2018) 97–107, ScienceDirect.

[2]. Raman Kumara and Gurpreet Singhb, “Analysis And Design Of An Optimized Secure Auditing Protocol For Storing Data Dynamically In Cloud Computing” International Conference on Processing of Materials, Minerals and Energy (July 29th – 30th) 2016, ScienceDirect, Ongole.

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

[3]. Reyhaneh Rabaninejad a, Mahmoud Ahmadian Attari a, Maryam Rajabzadeh Asaar b,∗, Mohammad Reza Aref c, "Comments on a lightweight cloud auditing scheme: Security analysis and improvement" Journal of Network and Computer Applications, Journal of Network and Computer Applications Volume 139, 1 August 2019, Pages 49-56.

[4]. Hao Jin, Hong Jiang, Senior Member, IEEE, and Ke Zhou, "Dynamic and Public Auditing with Fair Arbitration for Cloud Data" ,IEEE Transactions On Cloud Computing, Vol. 13, No. 9, September 2014.

[5]. Zheng Rujia ,Yu Ziya b , Wang Zhenkai c , "Remote audit for large data-based algebraic signature DCT cloud storage", ScienceDirect Measurement Volume 143, September 2019, Pages 22-26.

[6]. Hui Tian, , Yuxiang Chen, Chin-Chen Chang, , Hong Jiang Yongfeng Huang, Senior Member, IEEE, Yonghong Chen, "Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage" . IEEE Transactions On Sevice Computing, Manuscript Id 1 September 2015.

[7]. Anirudha Pratap Singha , Syam Kumar Pasupuletib , "Optimized Public Auditing and Data Dynamics for Data Storage Security in Cloud Computing" ,International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India,∗ Procedia Computer Science 93 (2016) 751 – 759.

[8]. Sookhak a,n , Hamid Talebian a , Ejaz Ahmed a , Abdullah Gani a , Muhammad Khurram Khan b, "A review on remote data auditing in single cloud server: Taxonomy and open issues",Mehdi Journal of Network and Computer Applications Volume 43, August 2014, Pages 121-141.

[9]. Geeta C Ma *, Raghavendra Sb , Rajkumar Buyyac , Venugopal K Rd , S S Iyengare , L M Patnaik, "Data Auditing and Security in Cloud Computing: Issues, Challenges and Future" Directions International Journal of Computer (IJC) ISSN 2307-4523 (Print & Online) © Global Society of Scientific Research and Researchers International Journal of Computer (IJC) (2018) Volume 28, No 1 , pp 8-57.

[10]. YanXu,SongSun,JieCui, HongZhong, "Intrusion-resilient public cloud auditing scheme with authenticator update" , Information Sciences Volume 512, February 2020, Pages 616-628.

[11]. Yu Fan1 , Yongjian Liao1 , (Member, Ieee), Fagen Li2 , (Member, Ieee), Shijie Zhou1 , Ganglin Zhang 1, "Identity-based Auditing for Shared Cloud Data with Efficient and Secure Sensitive Information Hiding" IEEE Access: DOI 10.1109/ACCESS.2019.2932430.

[12].Rafael Sanchez-Marquez a,∗, José Miguel Albarracín Guillemb, Eduardo Vicens-Salort a, José Jabaloyes Vivas, "Diagnosis of quality management systems using data analytics – A case study in the manufacturing sector", Computers in IndustryVolume 115, February 2020, 103183.

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

[13].System Pei Huang1 , Kai Fan1 , (Member, Ieee), Hanzhe Yang1 , Kuan Zhang2 , (Member, Ieee), Hui Li1 , (Member, Ieee), And Yintang Yang3 , "A Collaborative Auditing Blockchain for Trustworthy Data Integrity in Cloud Storage" Author et al.: Preparation of Papers for IEEE Transactions And Journals ,Volume 4, 2016.

[14].Mehdi Sookhak, Member, IEEE, F. Richard Yu, Senior Member, IEEE, and Albert Y. Zomaya, , "Auditing Big Data Storage in Cloud Computing Using Divide and Conquer Tables" IEEE Transactions On Parallel And Distributed Systems 1 1045-9219 (C) 2017.

[15].Yang Xu, Member, IEEE, Ju Ren, Member, IEEE, Yan Zhang, Senior Member, IEEE Cheng Zhang, Bo Shen, and Yaoxue Zhang, Senior Member, IEEE " Blockchain Empowered Arbitrable Data Auditing Scheme for Network Storage as a Service" , JOURNAL OF LATEX CLASS FILES, VOL. , NO. , MONTH 2019 .

[16].Jigang Wu∗,a , Guiyuan Jiangb , Thambipillai Srikanthanb a School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China , "Block chain-based public auditing for big data in cloud storage". Information Processing & Management Volume 57, Issue 6, November 2020, 102382.

[17]. Junfeng Tian, Haoning Wang, "A provably secure and public auditing protocol based on the bell triangle for cloud data" ,Computer Networks Volume 195, 4 August 2021, 108223.

[18].XiangGao[a]JiaYu[ab]Wen-TingShen[a]YanChang[c]Shi-BinZhang[c]MingYang[d]BinWu[b], "Achieving low-entropy secure cloud data auditing with file and authenticator DE duplication" Information Sciences science direct Volume 546, 6 February 2021, Pages 177-191.

[19].Neenu Garg,Seema Bawa,Neeraj Kumar "An efficient data integrity auditing protocol for cloud computing", Future Generation Computer Systems Volume 109, August 2020, Pages 306-316.

[20]. Jaya Rao Gudeme a,b,∗, Syamkumar Pasupuleti b, Ramesh Kandukuri a , "Certificateless privacy preserving public auditing for dynamic shared data with group user revocation in cloud storage", Journal of Parallel and Distributed Computing Volume 156, October 2021, Pages 163-175.

[21]. J. Wang, X. Chen, J. Li, J. Zhao, J. Shen, Towards achieving flexible and verifiable search for outsourced database in cloud computing, Future Gener. Comput. Syst. 67 (10) (2017) 266–275.

[22]. YuanXueaYu-anTanaChenLiangaYuanzhangLiaJunZhengaQuanxinZhang, "RootAgency: A digital signature-based root privilege management agency for cloud terminal devices" Information sciences Volume 444, May 2018, Pages 36-50

[23]. Rama Krishna Kalluri a , Guru C.V "An effective analytics of third party auditing and Trust architectures for integrity in cloud environment" Journal of King Saud University - Computer and Information Sciences2 March 2019

 https://doi.org/10.37896/pd91.4/91461

ISSN: 0369-8963

[24]. A.P. Singh, S.K. Pasupuleti, Optimized public auditing and data dynamics for data storage security in cloud computing, Procedia Comput. Sci. 93 (2016) 751

[25]. S. More, S. Chaudhari, Third party public auditing scheme for cloud storage, Procedia Comput. Sci. 79 (2016) 69–76.

[26]. J. Li, Y. Zhang, X. Chen, Y. Xiang, Secure attribute-based data sharing for resource-limited users in cloud computing, Comput. Secur. 72 (2018) 1–12, doi:10.1016/j.cose.2017.08.007.

 https://doi.org/10.37896/pd91.4/91461