

# Enabling Identity-based Integrity Auditing and Data Sharing with Sensitive Information Hiding for Secure Cloud Storage

Prof. C. P Lachake<sup>1</sup>, Shinde Balaji<sup>2</sup>, Ingle Shrikant<sup>3</sup>, Shinde Shradha<sup>4</sup>, Vedant Barapatre<sup>5</sup>

Professor, Department Of Computer Engineering<sup>1</sup>

Students, Department Of Computer Engineering<sup>2,3,4,5</sup>

SKN Sinhgad Institute of Technology and Science, Kusgaon (BK), Pune, Maharashtra, India

**Abstract:** A data storage server, such as a cloud server, can demonstrate to a verifier that it is honestly storing the data of a data owner by using remote data integrity checking (RDIC). Numerous RDIC protocols have been put out in the literature to this point, however the most of these designs have a sophisticated key management problem, meaning they depend on pricy public key infrastructure (PKI), which could make it difficult to implement RDIC in practise. In order to simplify the system and lower the cost of setting up and maintaining the public key authentication framework in PKI based RDIC schemes, we suggest a novel architecture of the identity-based (ID-based) RDIC protocol in this study. We formalise ID-based RDIC, along with its security model, which includes protection from rogue cloud servers and zero knowledge privacy from a third-party verification. During the RDIC procedure, the proposed ID-based RDIC protocol does not reveal any information about the stored data to the verifier. The new design achieves zero knowledge privacy against a verifier and is demonstrated to be secure against the malicious server in the general group model. Extensive security research and implementation results show that the suggested protocol is practicable in real-world applications and provably secure. We Extend This Work with Group Management, Forward and Backward Secrecy by Time Duration, and File Recovery When Data Integrity Checking Fault Occurs.

**Keywords:** Cloud Storage

## I. INTRODUCTION

A distributed computation model over a vast pool of shared-virtualized computing resources, such as storage, processing power, applications, and services, cloud computing has drawn a lot of interest from research communities in academia and industry. In a cloud computing environment, users can provide and release resources as they see fit. This new type of computing model is a fresh approach to offering computing services as essential public services like electricity and water. Many advantages come with cloud computing for users. However, a wide range of obstacles must be overcome before cloud computing may be widely used. Security concerns account for 87% of cloud users' worries, according to a recent Oracle survey that used data from the IDC Enterprise Panel1. One of the principal since customers no longer physically retain their data and lose control over it, their outsourced files' integrity is a security problem for cloud users. Furthermore, the cloud server cannot be completely trusted, and it is not required to notify instances of data loss. In fact, the cloud security alliance (CSA) presented an examination of cloud vulnerability occurrences to determine the dependability of cloud computing.

## II. RELATED WORK

The distributed file system at the far site ensures availability by copying each file to numerous desktop computers. It's critical to reclaim used space when you can because this replication uses a lot of storage capacity. Over 500 desktop file systems were measured, and the results reveal that duplicate files take up about half of the total space used. We outline a method for recovering space from this unintentional duplication so that it can be used for carefully regulated file replication. Our mechanism uses SALAD, a self-arranging, lossy, associative database, to aggregate file content and location data in a decentralised, scalable, fault-tolerant manner, as well as convergent encryption, which allows

duplicate files to merge into the space of a single file even if they are encrypted with different users' keys. Experiments with large-scale simulations reveal that the duplicate-file coalescing system is highly efficient, fault-tolerant, and scalable.

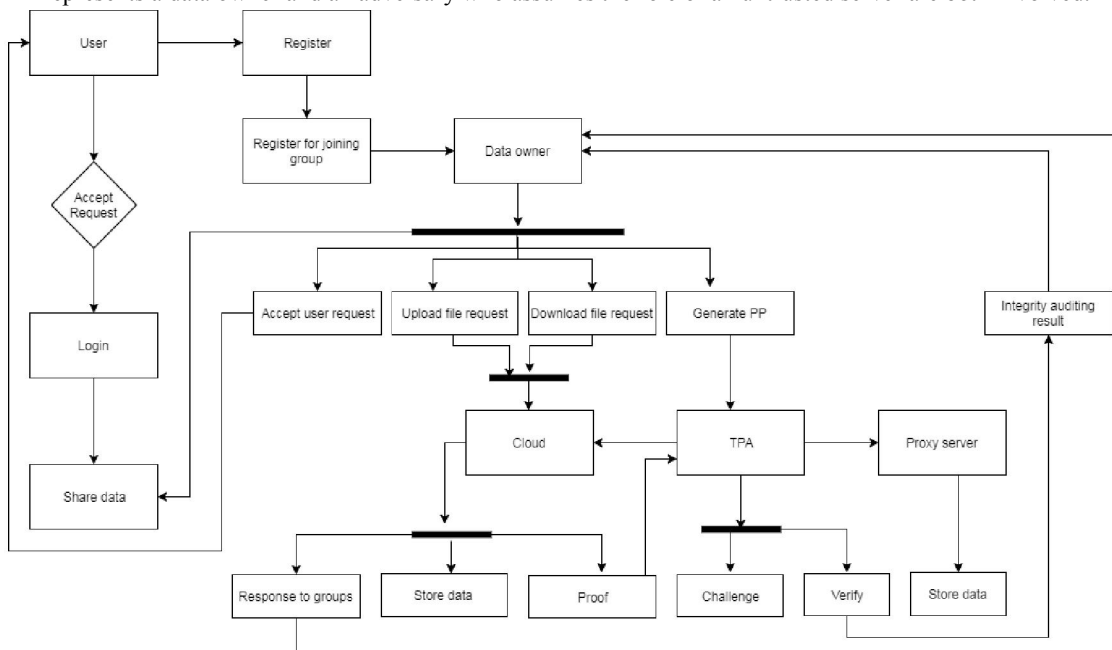
### III. SYSTEM ARCHITECTURE

Overall system design consists of following modules:

1. User
2. Cloud Sever
3. Third Party Auditor

In identity-based remote data integrity checking protocols, we take into account three security aspects, namely completeness, security against a malicious server (soundness), and privacy against the TPA (absolute data privacy). According to Shacham and Waters' security theories [7], an identity-based RDIC scheme is said to be secure against a server if neither a polynomial-time algorithm nor a polynomial-time extractor exists that can successfully circumvent the challenges-response protocols and recover the file. Completeness states that the ProofCheck algorithm will accept the proof when communicating with a legitimate cloud server. According to soundness, a cheating prover who is able to persuade the TPA that it is saving the data file is truly doing so. Now that the security is formalized

Model of soundness for identity-based remote data integrity checking is shown below, where a challenger who represents a data owner and an adversary who assumes the role of an untrusted server are both involved.



### IV. METHODOLOGY

Using key-homomorphic cryptographic primitives, we suggest a new architecture of the identity-based (ID-based) RDIC protocol in Fig. 1 to simplify the system and lower the cost of setting up and maintaining the public key authentication framework in PKI-based RDIC schemes.

We formalise ID-based RDIC, along with its security model, which includes protection from rogue cloud servers and zero knowledge privacy from a third-party verification.

During the RDIC procedure, the proposed ID-based RDIC protocol does not reveal any information about the stored data to the verifier.

#### **V. LITERATURE SURVEY**

Compare the relative merits of prevailing security theories for public key encryption systems. We take into account the objectives of privacy and non-malleability, separately, for two different types of chosen ciphertext attacks and a chosen plaintext assault. We demonstrate either an implication or a contradiction for each of the resulting pairs of definitions. We describe a fresh method for swiftly revoking user credentials and fine-grained control over their security rights that is based on the idea of an online semi-trusted mediator (SEM). There are a number of practical benefits to using a SEM in conjunction with the mediated RSA cryptosystem as opposed to existing revocation methods. The advantages include faster revocation of signature and decryption capabilities as well as streamlined certificate revocation for legacy systems. "public key" systems allowing data encryption queries Tokens can be created using a secret key to test any supported query predicate. Without learning anything else about the plaintext, the token enables anyone to test the predicate on a given cypher text.

#### **VI. FUTURE SCOPE**

We provide evidence that the suggested plan successfully satisfies the criteria for completeness, soundness, and perfect data privacy preservation. While soundness demonstrates that the protocol is secure against an untrusted server, completeness ensures that the protocol is proper. The definition of perfect data privacy is the protocol not disclosing any information about the stored files to the verifier.

#### **VII. CONCLUSION**

In this, we looked into a brand-new basic for safe cloud storage called identity-based remote data integrity checking. We defined the security model of this primitive's two key characteristics, soundness and absolute data privacy. We gave this primitive a fresh design and demonstrated how it achieves soundness and full data privacy. The proposed protocol's effectiveness and applicability were demonstrated by both the numerical analysis and the execution. Group Management with Forward Secrecy and Backward Secrecy by Time Duration & File Recovery when Data Integrity Checking Fault Occurs are extensions of this work.

#### **REFERENCES**

- [1]. P. Mell, T. Grance, Draft NIST working definition of cloud computing, Reference on June. 3rd, 2009. <http://csrc.nist.gov/groups/SNC/cloudcomputing/index.html>.
- [2]. Cloud Security Alliance. Top threats to cloud computing. <http://www.cloudsecurityalliance.org>, 2010.
- [3]. M. Blum, W. Evans, P. Gemmell, S. Kannan, M. Naor, Checking the correctness of memories. Proc. of the 32nd Annual Symposium on Foundations of Computers, SFCS 1991, pp. 90–99, 1991.
- [4]. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner, Z. N.J. Peterson, D. X. Song, Provable data possession at untrusted stores. ACM Conference on Computer and communications Security, 598-609, 2007.
- [5]. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. N. J. Peterson, and D. Song, Remote data checking using provable data possession. ACM Trans. Inf. Syst. Secur., 14, 1–34, 2011.
- [6]. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. N. J. Peterson, and D. Song, Remote data checking using provable data possession. ACM Trans. Inf. Syst. Secur., 14, 1–34, 2011.
- [7]. A. Juels, and B. S. K. Jr. Pors, proofs of retrievability for large files. Proc. of CCS 2007, 584-597, 2007.
- [8]. H. Shechem, and B. Waters, Compact proofs of retrievability. Proc. of Cryptology-ASIACRYPT 2008, LNCS 5350, pp. 90-107, 2008.
- [9]. G. Ateniese, S. Kamara, J. Katz, Proofs of storage from homomorphic identification protocols. Proc. of ASIACRYPT 2009, 319-333, 2009.
- [10]. A. F. Barsoum, M. A. Hasan, Provable multicopy dynamic data possession in cloud computing systems, IEEE Trans. on Information Forensics and Security, 10(3): 485–497, 2015