International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 23 Issue 8 –NOVEMBER 2016.

Dynamic Data Possession and Outsourcing Data Storage with Provable Multicopy in Cloud

MOHAMMAD JAINAB MOTHI[#1], GUNTAPALLI MINNI[*2]and SAYEED YASIN[*3]

[#] *M.Tech (CSE) Student, Nimra College of Engineering & Technology, A.P., India.*

[*] *Assistant Professor, Dept. of Computer Science & Engineering, Nimra College of Engineering & Technology, A.P., India.*

[*]*Associate professor & Head, Dept. of Computer Science & Engineering, Nimra College of Engineering & Technology, A.P., India.*

***Abstract*— Cloud Computing (CC) is an emerging computing paradigm that can offer multiple advancements. Outsourcing data to a remote cloud service provider (CSP) allows organizations to store more data on the CSP than on private computer systems. Such outsourcing of data to CSP helps organizations from computing issues and to concentrate more on innovations. All the authorized users can retrieve data from different locations in the world from CSP. The data owners lose the direct control over their sensitive data once the data has been outsourced to a remote CSP which need not be trustworthy. This lack of control over their sensitive data raises new ominous and challenging tasks related to data confidentiality and integrity maintenance in cloud computing. The confidentiality issue can be solved by encrypting the data before outsourcing to cloud storage. As such, it is a high priority demand of customers to have strong evidence that the cloud servers still possess their data and it is not being tampered with or partly deleted over time. Accordingly, many researchers have focused on the problem of Provable Data Possession (PDP) and proposed different techniques to review the data over remote servers. PDP is a technique to approve data integrity over remote servers. So, propose system introduced the multi owner with data sharing concept. Most of the time customers store their important data on multiple servers for accessing on different geographical location. When data store on multiple server then data owner need evidence of all outsourced file copies are intact. But some time few copies may be corrupt. So, proposed system identifies corrupted copies and corrects it before dynamic operation performs. In addition, proposed system allows multiowner facility for sharing data.**

Index Terms*— *Cloud computing, Cloud Service Provider, Data owner, Data Integrity, Multi-copy.

I. INTRODUCTION

Cloud computing provides shared processing environment for data storage and accessing also known as internetbased computing. It is a model which provides configurable computing resources such as networks, servers, storage, applications and services. Cloud computing has a high computation power, lowest cost of services, higher performance, scalability, accessibility and availability for that reason it is highly demanded. Data outsourcing brings with it many advantages. But associated with it are the risks involved. Though client cannot physically access the data from the cloud server directly, without client's are either not used by client from a long time. Hence, there is a requirement of checking the data periodically for correction purpose, known as data integrity. Here we provide a survey on the different techniques of data integrity. The basic schemes for data integrity in cloud are Provable Data Possession (PDP) knowledge, cloud provider can modify or delete data which and Proof of Retrievability (PoR). These two schemes are the most active area of research in the cloud data integrity field. In data integrity checking, client challenge remote server and server response by proving that. Many researchers have focused on this problem and find out different technics. PDP is one of the techniques for validating data integrity. In this model, do not need to store all file to local computer to check data integrity. It creates metadata information of each file and that store it in local computer without storing whole file. At the time of verification of data integrity it sends the metadata to the verifier side. PDP model used both static data and dynamic data. In static PDP schema used data which cannot change, it only store and access by authorize users [2][3]. In dynamic PDP schema stored data can be modify by performing operation like modify, insert, delete etc. and also it can be scaled by inserting more data [4] [5] [6]. For efficient validation of outsourced data integrity, a number of PDP schema proposed, which is based on single copy [7] where no need to proof that CSP store all copies but in multiple copy CSP need to proof. In multiple data copies, the overall system integrity fails if there are one or more corrupted copies. Also, we need to check integrity each time when we performing dynamic operation. MB-PDP schema used dynamic data to store multiple copies on different server across different data centers [8]. In this proposed system, address all of this issue and recognize list of corrupted copies and reconstruct them. In addition to improve scalability, this system provides Multi-owner and sharing facility. An increasing number of clients organization use cloud to store data which has become trend [1]. Cloud service provider (CSP) allows storing much more data than private computer. Once data stored in remote

server, authorizer can access all data from any geographical location. Most of the time organization store important data in cloud, without leaving a copy in local computer. Once data is stored in cloud they may not be trustworthy due to losing control on data. So, it is important to ensure data is not lost or corrupted by checking data integrity.

II. PROPOSED SYSTEM

Existing system provides adequate guarantee that the CSP store all the copies of data that are agreed upon in contract. But it does not provide any proof whether the copies of the data are stored in different locations over different servers. In the proposed scheme data owners divide the file into blocks then it generate keys needed for the sessions. Then the blocks are encrypted and it is outsourced to the CSP. CSP receives the file and generate the copies of the as per in the contract. Then it send the file copies to respected locations and CSP creates location tags with location details. These location tags are shared with the data owner. Data owner send a challenge to the CSP for getting the proof. This can be done by an interactive zero knowledge protocol. Interactive zero knowledge protocol is a method in which one party can covey to the other party that the given statement is true. Non interactive protocol different class of Zero-Knowledge proof systems, where no interaction is required: The Prover simply sends one message to the Verifier, and the Verifier either accepts or rejects. Interactive zero knowledge protocol is used at the time of verification. From each location a proof is generated and is send to the data owner to verify. Verifier verifies the proof. CSP maintains insertion, deletion, modification of the blocks in the CSP. Authorised users can access the file by getting a shared key from the Data owner.

When verifying multiple data copies, the overall system integrity check fails if there is one or more corrupted copies. To address this issue and recognize which copies have been corrupted, we discuss a slight modification to be applied to the proposed scheme.

We propose a map-based provable multi-copy dynamic data possession (MB-PMDDP) scheme. This scheme provides an adequate guarantee that the CSP stores all copies that are agreed upon in the service contract. Moreover, the scheme supports outsourcing of dynamic data, *i.e.*, it supports block-level operations such as block modification, insertion, deletion, and append. The authorized users, who have the right to access the owner's file, can seamlessly access the copies received from the CSP.

We give a thorough comparison of MB-PMDDP with a reference scheme, which one can obtain by extending existing PDP models for dynamic single-copy data.

We show the security of our scheme against colluding servers, and discuss a slight modification of the proposed scheme to identify corrupted copies.

In this work, we do not encode the data to be outsourced for the following reasons. First, we are dealing with dynamic data, and hence if the data file is encoded before outsourcing, modifying a portion of the file requires re-encoding the data file which may not be acceptable in practical applications due to high computation overhead.

Second, we are considering economically-motivated CSPs that may attempt to use less storage than required by the service contract through deletion of a few copies of the file. The CSPs have almost no financial benefit by deleting only a small portion of a copy of the file.

Third, and more importantly, unlike erasure codes, duplicating data files across multiple servers achieves scalability which is a fundamental customer requirement in CC systems. A file that is duplicated and stored strategically on multiple servers – located at various geographic locations

A. ADVANTAGES OF PROPOSED SYSTEM:

Although PDP schemes have been presented for *multiple* copies of *static* data, to the best of our knowledge, this work is the first PDP scheme directly dealing with *multiple* copies of *dynamic* data.

The storage model used in this work can be adopted by many practical applications. For example, e-Health applications can be envisioned by this model where the patients' database that contains large and sensitive information can be stored on the cloud servers. In these types of applications, the e-Health organization can be considered as the data owner, and the physicians as the authorized users who have the right to access the patients' medical history. Many other practical applications like financial, scientific, and educational applications can be viewed in similar settings.

III. LITERATURE SURVEY

#1 Demonstrating data possession and uncheatable data transfer

We observe that a certain RSA-based secure hash function is homomorphic. We describe a protocol based on this hash function which prevents 'cheating' in a data transfer transaction, while placing little burden on the trusted third party that oversees the protocol. We also describe a cryptographic protocol based on similar principles, through which a prover can demonstrate possession of an arbitrary set of data known to the verifier. The verifier isn't required to have this data at hand during the protocol execution, but rather only a small hash of it. The protocol is also provably as secure as integer factoring.

A. Disadvantages

1. Less security due to single cloud
2. There is no Reconstruction method to overcome the attacked data retrieval.

#2 Efficient Remote Data Possession Checking In Critical Information Infrastructures Ensuring Data Storage Security In Cloud Computing

Cloud computing has been envisioned as the on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing and transference of risk. Today, technical research

works focus on Remote data possession Checking protocols permit to check that a remote server can access an uncorrupted file with the help of third party verifiers. In this paper, Seb´e et al.'s protocol is adapted to support efficient remote data possession checking in critical information infrastructure without the help of a third party auditor. This design allows users to audit the cloud storage with very lightweight communication and computation cost.

In addition, the auditing result not only ensures strong cloud storage correctness guarantee, but also simultaneously achieves fast data error localization, i.e., the identification of misbehaving remote server. The design further supports secure and efficient dynamic operations on outsourced data, including block modification, deletion, and append. Through a formal analysis, the correctness and security of the protocol is shown. The proposed scheme is highly efficient and resilient against the malicious data modification attack, server clouding attacks and failure.

B. Disadvantages

1. There is no feature of automatic blocking the cloud server attackers.
2. Owner data will be stored in untrusted cloud servers.

#3 Auditing to Keep Online Storage Services Honest

A growing number of online service provider's offers to store customers' photos, email, file system backups, and other digital assets. Currently, customers cannot make informed decisions about the risk of losing data stored with any particular service provider, reducing their incentive to rely on these services. We argue that third party *auditing* is important in creating an online service oriented economy, because it allows customers to evaluate risks, and it increases the efficiency of insurance based risk mitigation. We describe approaches and system hooks that support both *internal* and *external* auditing of online storage services, describe motivations for service providers and auditors to adopt these approaches, and list challenges that need to be resolved for such auditing to become a reality.

C. Disadvantages

1. There is no feature of automatic blocking the cloud server attackers.
2. Less Security – No cryptographic technique is used on the cloud data

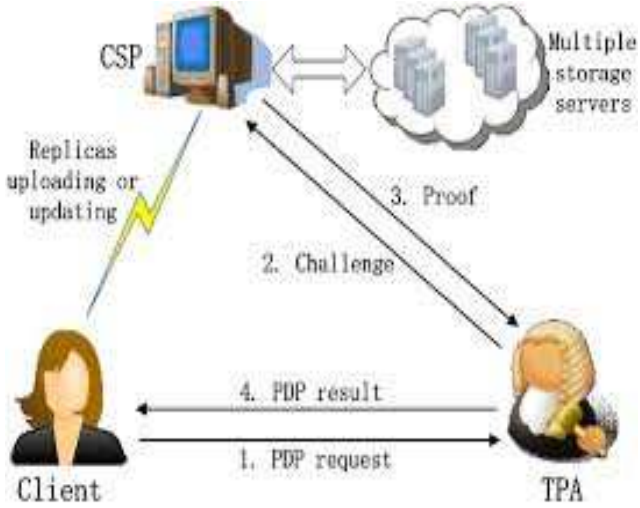
#4 Authentications and Integrity in Outsourced Databases

In the Outsourced Database (ODB) model, entities outsource their data management needs to third-party service providers. Such a service provider offers mechanisms for its clients to create; store, update and access (query) their databases. This work provides mechanisms to ensure data integrity and authenticity for outsourced databases. Specifically, this work provides mechanisms that assure the querier that the query results have not been tampered with and is authentic (with respect to the actual data owner). It investigates both security and efficiency aspects of the problem and constructs several secure and practical schemes that facilitate integrity and authenticity of query replies while incurring low computational and communication costs.

D. Disadvantages

1. The data integrity is proving only based on the filename and not on the public key or any other key.

3. The attackers details are not dynamic instead its maintaining the log files to store the attacker details and viewing using data mining concepts which is time consuming job and less security.



IV. REALATED WORK

Researchers have proposed two basic approaches to client verification of file availability and integrity. The cryptographic community has proposed tools called proofs of retrievability (PORs) and proofs of data possession (PDPs).PDP scheme checks that a remote cloud server retains a file, which consists of a collection of n blocks. The data owner processes the data file to generate some metadata to store it locally. The file is then sent to the server, and the owner delete the local copy of the file. The owner verifies the possession of file in a challenge response protocol. A POR is a challenge response protocol that enables a prover (cloud-storage provider) to demonstrate to a verifier (client) that a file F is retrievable, i.e., recoverable without any loss or corruption. The benefit of a POR over simple transmission of F is efficiency. The response can be highly compact (tens of bytes), and the verifier can complete the proof using a small fraction of F.As a standalone tool for testing file retrievability against a single server, though, a POR is of limited value. Detecting that a file is corrupted is not helpful if the file is irretrievable and the client has no recourse. Thus PORs are mainly useful in environments where F is distributed across multiple systems, such as independent storage services. In

such environments, F is stored in redundant form across multiple servers.

i) **Data Owner:-** Data owner generates keys that are required for sessions.

- It divides the files into blocks.
- These blocks are encrypted.
- These blocks are outsourced to the CSP.
- It receives location tags from the CSP and maintains the location details in it.
- It challenges the CSP to provide proof. It sends challenge to the CSP to verify whether the agreed number of copies

are stored in the CSP.

Proof is received by the data owner from different locations that are specified in the tags.

- After receiving the proof, proof is verified. If proof is correct then the exact copies of the files are maintained in the CSP. Then the data owner confirms reliability with the CSP.

When the authorized users request to grant permission to access the file, data owner will share a key with the user and user will access the file with it.

ii) **Cloud Service Provider:-** CSP receives the file blocks outsourced to it.

- CSP creates multiple copies that agreed with the data owner.
- It sends the file copies to the location.
- After sending the file to location, tags are created with the details of the location.
- These created location tags are send to data owner.
- CSP receives a challenge from the data owner.
- When challenge is received, it is passed to the locations where the copies are stored.
- Each location computes a proof and these proofs are passed to the data owner with interactive zero knowledge

protocol. Operations like insertion, deletion, modification, append are performed in the CSP on file blocks according to data owners' request. Insertion insert a block anywhere in the file. Deletion deletes the block completely. Modification modifies the block content. Append operation adds a new block at the end of the blocks. After the operation change must be updated to all the copies present in the CSP.

- Request for accessing the file is received from the authorized users.
- After checking the authenticity encrypted blocks are send to the authorized users.

iii) **Authorised Users:-** Authorized users request the data owner to grant permission to access the file from the CSP.

- It will receive a key from the data owner.
- After receiving the key, it will request for the file to the CSP.
- User will receive the encrypted blocks of the file in an unordered manner.
- Blocks are decrypted using the Shared secret key. These blocks are rearranged to get a complete file. Every file can be decrypted with the same key. Users can seamlessly access the file from the CSP.

V. CONCLUSION

Outsourcing data to remote servers has become a growing trend for many organizations to Alleviate the burden of local data storage and maintenance. In this work we have studied the problem of creating multiple copies of dynamic data file and verifying those copies stored on untrusted cloud servers. We have proposed a new PDP scheme (referred to as MB-PMDDP), which supports outsourcing of multi-copy dynamic data, where the data owner is capable of not only archiving and accessing the data copies stored by the CSP, but also updating and scaling these copies on the remote servers. To the best of our knowledge, the proposed scheme is the first to address *multiple* copies of *dynamic* data. The interaction between the authorized users and the CSP is considered in our scheme, where the authorized users can seamlessly access a data copy received from the CSP using a single secret key shared with the data owner. Moreover, the proposed scheme supports public verifiability, enables arbitrary number of auditing, and allows *possession-free* verification where the verifier has the ability to verify the data integrity even though he neither possesses nor retrieves the file blocks from the server. Through performance analysis and experimental results, we have demonstrated that the proposed MB-PMDDP scheme outperforms the TB-PMDDP approach derived from a class of dynamic single-copy PDP models. The TB-PMDDP leads to high storage overhead on the remote servers and high computations on both the CSP and the verifier sides. The MB-PMDDP scheme significantly reduces the computation time during the challenge-response phase which makes it more practical for applications where a large number of verifiers are connected to the CSP causing a huge computation overhead on the servers. Besides, it has lower storage overhead on the CSP, and thus reduces the fees paid by the cloud customers. The dynamic block operations of the map-based approach are done with less communication cost than that of the tree-based approach. A slight modification can be done on the proposed scheme to support the feature of identifying the indices of corrupted copies. The corrupted data copy can be reconstructed even from a complete damage using duplicated copies on other servers. Through security analysis, we have shown that the proposed scheme is provably secure.

REFERENCES

[1] G. Ateniese et al., "Provable data possession at untrusted stores," in Proc. 14th ACM Conf. Comput. Commun. Secur. (CCS), New York, NY,USA, 2007, pp. 598–609.

[2] K. Zeng, "Publicly verifiable remote data integrity," in Proc. 10th Int. Conf. Inf. Commun. Secur. (ICICS), 2008, pp. 419–434.

[3] Y. Deswarte, J.-J. Quisquater, and A. Saïdane, "Remote integrity checking," in Proc. 6th Working Conf. Integr. Internal Control Inf. Syst. (IICIS), 2003, pp. 1–11.

[4] D. L. G. Filho and P. S. L. M. Barreto, "Demonstrating data possession and uncheatable data transfer," IACR (International Association for Cryptologic Research) ePrint Archive, Tech. Rep. 2006/150, 2006.

[5] F. Sebé, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient remote data possession checking in critical information infrastructures," IEEE Trans. Knowl. Data Eng., vol. 20, no. 8, pp. 1034–1038, Aug. 2008.

[6] P. Golle, S. Jarecki, and I. Mironov, "Cryptographic primitives enforcing communication and storage complexity," in Proc. 6th Int.

Conf. Financial Cryptograph. (FC), Berlin, Germany, 2003, pp. 120–135.

[7] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in Proc. 11th USENIX Workshop Hot Topics Oper. Syst. (HOTOS), Berkeley, CA, USA, 2007, pp. 1–6.

[8] M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," IACR Cryptology ePrint Archive, Tech. Rep. 2008/186, 2008.

[9] E. Mykletun, M. Narasimha, and G. Tsudik, "Authentication and integrity in outsourced databases," ACM Trans. Storage, vol. 2, no. 2, pp. 107–138, 2006.

[10] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in Proc. 4th Int. Conf. Secur. Privacy Commun. Netw. (SecureComm), New York, NY, USA, 2008, Art. ID 9.

[11] C. Wang, Q. Wang, K. Ren, and W. Lou. (2009). "Ensuring data storage security in cloud computing," IACR Cryptology ePrint Archive, Tech. Rep. 2009/081. [Online]. Available: http://eprint.iacr.org/

[12] C. Erway, A. Küpçü, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," in Proc. 16th ACM Conf. Comput. Commun. Secur. (CCS), New York, NY, USA, 2009, pp. 213–222.

MOHAMMAD JAINAB MOTHI is a student of Nimra college of engineering and Technology, Jupudi, NimraNagar, VIJAYAWADA. She is presently pursuing her M.Tech degree from JNTU, Kakinada.She has obtained B.Tech degree from JNTU, Kakinada.

G.MINNI is presently working as Assistant professor in CSE department in Nimra college of Engineering and Technology, Jupudi, Nimra Nagar,VIJAYAWADA. She has obtained M.Tech degree from JNTU, Kakinada. She is pursuing Ph.D., in A.N.U, GUNTUR. She has published several research papers in various national and international Journals. She has more than Ten years of experience in teaching field, her area of interests are networks & Web Designing.

SAYEED YASIN received his M.TECH in Computer Science & Engg from JNTU Hyderabad. He is pursuing Ph.D., in Rayalaseema University, Kurnool. He is currently working as an Associate Professor & Head in Nimra College of Science & Technology the Department of Computers Science and Engineering & Technology, Jupudi, Ibrahimpatnam,Vijayawada-521456. He has more than Eight years of experience in teaching field. His area of interests are wireless networks & programming, & Mobile Computing.