ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Distributed Provable Data Possession and Data Integrity Verification in Multi-Cloud Storage Using Client Authentication

Avinash Sanjay Pawar
M.Tech Student,
Department of CSE,
Marri Laxman Reddy Institute of Technology and Management.

Dr.K.Venkateswara Reddy
Professor,
Department of CSE,
Marri Laxman Reddy Institute of Technology and Management.

Abstract:

Generally the cloud server act as a container which contains data or information. Providing Security in cloud is becoming a difficult task now days. Remote data integrity checking is of crucial importance in cloud storage. It can make the clients verify whether their outsourced data is kept intact without downloading the whole data. In some application scenarios, the clients have to store their data on multi-cloud servers. At the same time, the integrity checking protocol must be efficient in order to save the verifier's cost. From the two points, we propose a novel remote data integrity checking model: ID-DPDP (identity-based distributed provable data possession) in multi-cloud storage. In addition to the structural advantage of elimination of certificate management, our ID-DPDP protocol is also efficient and flexible.

Keywords:
Cloud computing, Provable data possession, Identity-based cryptography, Distributed computing.

INTRODUCTION:

The significance of cloud computing has made it important in the field of computers by consideration of storage, computing and so on. The basis of cloud system lies within computing tasks outsourcing towards third party and necessitates security risks concerning reliability as well as accessibility of data and service. In cloud system, security problem of remote data integrity verification is an important issue considered [1]. For gaining of additional benefits, malicious cloud server might damage data of client data. The concept of provable data possession was introduced by Ateniese et al in which verifier verifies data integrity by means of a high probability. On the basis of distributed computation, we make a learning of distributed model of data integrity checking and provide a concrete protocol within multi-cloud storage.

We introduce new data integrity checking representation: known as identity-basis distributed provable data possession in the systems of multi-cloud storage. For improving of the efficiency, there is a consideration of identity-based provable data possession to be more attractive and hence it is significant for making a study of identity-basis distributed provable data possession. By considering the structural benefit of eliminating certificate management, our proposed system of identity-based provable data possession is moreover flexible. On the basis of provable data possession procedure, system of identity-basis distributed provable data possession is constructed by usage of signature as well as distributed computing. cloud to be subject for any information debasement or security rupture into which their cloud administration supplier (CSP) may bring about, actually when they don't hold control over their information. Persuading cloud clients that their information is in place is particularly indispensable when clients are organizations. Remote information ownership checking (RDPC) is a primitive intended to address this issue. Ateniese et al. [2] have formalized a model called provable data ownership (PDP). In this model, information (regularly spoke to as an issue F) is preprocessed by the customer, and metadata utilized for check purposes is delivered. The document is then sent to an untrusted server for capacity, and the customer may erase the neighborhood duplicate of the record. The customer keeps some (conceivably mystery) data to check server's reactions later. The server demonstrates the information has not been messed around with by reacting to difficulties sent by the customer. The creators display a few varieties of their plan under distinctive cryptographic suspicions. These plans give probabilistic insurances of ownership, where the customer checks an irregular subset of put away hinders with each one test. Notwithstanding, PDP and related plans [2, 7, 12] apply just to the instance of static, archival stockpiling, i.e., a record that is outsourced and never shows signs of change (at the same time with our work,

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Ateniese et al. [3] present a plan with sort of restricted dynamism, which is talked about in subtle element in the related work segment). While the static model fits some application situations (e.g., libraries and experimental datasets), it is significant to consider the element case, where the customer upgrades the outsourced information by embedding, changing, or erasing put away pieces or documents while keeping up information ownership ensures. Such an element PDP plan is fundamental in commonsense distributed computing frameworks for document stockpiling [13, 16], database administrations [17], and shared capacity [14, 20]. The data owners lose the management over their sensitive data once the latter is outsourced to a distant CSP which cannot be trustworthy. This lack of management raises new formidable and difficult tasks related to data confidentiality and integrity protection in cloud computing. Customers need that their data stay secure over the CSP. Also, they have to possess a robust proof that the cloud servers still possess the data and it is not being tampered with or partly deleted over time, particularly as a result of the internal operation details of the CSP might not be notable to cloud customers. Encrypting sensitive data before outsourcing to remote servers will handle the confidentiality issue. However, the integrity of customers' knowledge within the cloud is also in danger due to the subsequent reasons. Several researchers have centered on provable of demonstrable data possession (PDP) and planned completely different schemes to audit the data validating on remote servers. PDP is a technique for validating data integrity over remote servers. In an exceedingly typical PDP model, the data owner generates some metadata/information for an information file to be used later for verification purposes through a challenge response protocol with the remote cloud server. The owner sends the file to be stored on a foreign server which can be untrusted, and deletes the native copy of the file. In PKI (public key infrastructure), obvious knowledge possession protocol desires public key certificate distribution and management. It will incur sizeable overheads since the verifier can check the certificate once it checks the remote data integrity. Additionally to the significant certificate verification, the system conjointly suffers from the opposite sophisticated certificates management reminiscent of certificates generation, delivery, revocation, renewals, etc. In cloud computing, most verifiers only have low computation capability. Identity-based public key cryptography will eliminate the sophisticated certificate management. So as to extend the potency, identity-based provable knowledge possession is a lot of enticing. Thus, it will be very significant to check the ID-RDP.

Motivation:

We consider an ocean information service corporation Cor in the cloud computing environment. Cor can provide the following services: ocean measurement data, ocean environment monitoring data, hydrological data, marine biological data, GIS information, etc. Besides of the above services, Cor has also some private information and some public information, such as the corporation's advertisement. Cor will store these different ocean data on multiple cloud servers. Different cloud service providers have different reputation and charging standard. Of course, these cloud service providers need different charges according to the different security-levels. Usually, more secure and more expensive. Thus, Cor will select different cloud service providers to store its different data. For some sensitive ocean data, it will copy these data many times and store these copies on different cloud servers. For the private data, it will store them on the private cloud server. For the public advertisement data, it will store them on the cheap public cloud server. At last, Cor stores its whole data on the different cloud servers according to their importance and sensitivity. Of course, the storage selection will take account into the Cor's profits and losses. Thus, the distributed cloud storage is indispensable. In multi-cloud environment, distributed provable data possession is an important element to secure the remote data. In PKI (public key infrastructure), provable data possession protocol needs public key certificate distribution and management. It will incur considerable overheads since the verifier will check the certificate when it checks the remote data integrity. In addition to the heavy certificate verification, the system also suffers from the other complicated certificates management such as certificates generation, delivery, revocation, renewals, etc. In cloud computing, most verifiers only have low computation capacity. Identity-based public key cryptography can eliminate the complicated certificate management. In order to increase the efficiency, identity-based provable data possession is more attractive. Thus, it will be very meaningful to study the ID-DPDP.

LITERATURE SURVEY:

In [1] authors introduced a model for provable data possession (PDP) that allows a client that has outsourced data at an untrusted cloud to verify that the server possesses the original data without downloading it. This model generates a probabilistic proof of possession through sampling

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

random set of blocks from the server, which significantly reduces cost. The data owner maintains a constant amount of data to verify the proof. The request/response protocol transmits a little, constant amount of data, which reduces network communication. Thus, the Provable Data Possession model for remote data integrity checking supports the large data sets in widely-distributed storage system. The key component of this scheme is the homomorphism verifiable tags. In [2] authors introduced the efficient and secured outsourced data is addressed either by public key cryptography or requiring the user to outsource its data in encrypted form called EPDP (Efficient-PDP). This technique is based entirely on symmetric key cryptography and not requiring any bulk encryption. It allows dynamic data that efficiently support operations, such as block modification, deletion and append.

Two different approaches PDP and POR have been proposed. The POR is a public key based technique allowing any verifier to query the server and obtain an interactive proof of data possession. In [3] authors proposed the POR scheme permits back-up service to produces a concise proof that a client can retrieve a file F, that is, that the archive retains and reliably transmits file data sufficient for the user to recover F in its whole. A POR is a kind of cryptographic proof of knowledge (POK), but one specially designed to handle a large file F. To explore POR protocols, in which the communication expenses, memory accesses for the proven, and storage necessities of the client are small parameters essentially independent of the length of F. The goal of a POR is to accomplish these checks without client having to retrieve the files themselves. A POR can also provide service with quality assurances.In [4] authors introduced the problem of ensuring the integrity of data storage.

In particular, to consider the task of allowing a third party auditor (TPA), on behalf of the user, to verify the integrity of the dynamic data stored in the cloud server. The introduction of TPA reduces the participation of the client through the auditing of whether their data in the cloud is indeed intact, which can be important in achieving financial system of scale for Cloud Computing. The operation supported by data dynamics such as block insertion and deletion, is also a major step toward practicality, services present in Cloud Computing are not limited to archive or backing data only. Earlier works on make confident of the remote data integrity often lacks the support of either public audit ability or dynamic data operations, this paper achieves both.

Initially identify the complexities and security problems of direct extensions with fully dynamic data operations from prior works and then show how to construct an elegant verification scheme for the seamless integration of these two salient features in this protocol design.In [6] authors considered the cloud data storage protection, which has always been an essential aspect of ensures the accuracy of client data in the cloud, it is denoting ineffective and flexible distributed verification scheme with two features. By utilizing the homomorphism token with flexible distributed verification achieves the storage correctness and data error localization. Unlike the most prior works, this scheme further supports secured and efficient dynamic operation son data blocks, including: data insert, update, delete and append. If supposed to find fraud in our outsourced data (e.g., when a server crashes or is compromised) in the storage cloud, then we should correct the corrupted data and restore the original data in the cloud.

In [7] authors improved the Remote data integrity checking can make the client to verify their outsourced data is kept intact without retrieving the entire data. In some application scenarios, the users have to store their data on multi-cloud environment. At the same time, the integrity checking protocol must be efficient in order to save the verifier's cost. From the two points, propose a novel remote data integrity checking model: IDDPDP (identity-based distributed provable data possession) in multi-cloud storage. Concrete ID-DPDP protocol is designed based on the bilinear pairing. The ID-DPDP protocol is provably secure under the hardness assumption of the standardCDH(computational DiffieHellman) problem. ID-DPDP protocol is also efficient and flexible because it eliminates the certificate management. Based on the client's authentication, the ID-DPDP protocol can realize private verification, delegated verification and open verification.

In [8] authors proposed Provable data possession (PDP) is a method for ensuring the integrity of data in cloud. In this paper, to address the construction of an efficient PDP scheme for distributed cloud storage to support the scalability of service and data migration, in which to consider the presence of multiple cloud service providers to cooperatively store and maintain the clients' data. To present a cooperative PDP (CPDP) scheme based on homomorphism verifiable reaction and hash index hierarchy. Prove the security of the scheme based on multi-proverb zero-knowledge proof system, which can fulfill completeness, information soundness, and zero-knowledge goods.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

In addition, articulate performance optimization mechanisms for this scheme, and in particular present an effective method for selecting optimal parameter values to minimize the computation costs of clients and storage service providers. This experiments show that our solution presents lesser computation and communication expenses in comparison with non-cooperative approaches. Cooperative PDP (CPDP) schemes accepting zero-knowledge property and three-layered index hierarchy, respectively. In particular effective method for choosing the optimal amount of sectors in each block to minimize the computation charges of clients and storage service providers. Cooperative PDP (CPDP) scheme without compromising data secrecy based on current cryptographic techniques. In [9] authors proposed about cloud storage, users can remotely store their client data and appreciate the on-demand high-quality presentations and services from a shared pool of configurable computing assets, without the load of native data storage and protection. The statement that client no larger have physical ownership of the uploaded data makes the data integrity security in cloud computing a difficult task, especially for clients with controlled computing assets. Moreover, users should be able to use the cloud storage as if it is resident, without distressing about the need to check its integrity. Thus, permitting open audit ability for cloud storage is of serious importance so that client can resort to a third-party auditor to verify the integrity of outsourced data and free. To strongly introduce an effective TPA, the checking process should bring in no new weaknesses toward client data privacy, and present no supplementary online burden to user. In this paper, to introduce a secure cloud storage structure supporting privacy-preserving public auditing. To further range this result to enable the TPA to perform verify for multiple users concurrently and powerfully. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient.

PROBLEM STATEMENT:

In cloud computing, remote data integrity checking is an important security problem. The clients' massive data is outside his control. The malicious cloud server may corrupt the clients' data in order to gain more benefits. The formal system model and security model are existing models. In the PDP model, the verifier can check remote data integrity with a high probability. Based on the RSA, they designed two provably secure PDP schemes. PDP allows a verifier to verify the remote data integrity without retrieving or downloading the whole data.

It is a probabilistic proof of possession by sampling random set of blocks from the server, which drastically reduces I/O costs. The verifier only maintains small metadata to perform the integrity checking. PDP is an interesting remote data integrity checking model. In POR, the verifier can check the remote data integrity and retrieve the remote data at any time. On some cases, the client may delegate the remote data integrity checking task to the third party. It results in the third party auditing in cloud computing.

DISADVANTAGES:

•Does not provide efficiency in remote data integrity checking.
•More expensive.
•The existing system provides less flexibility.

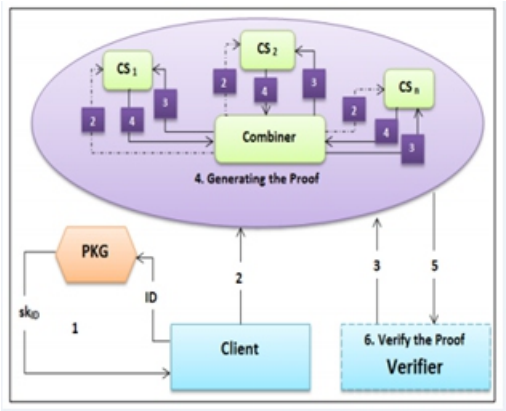
PROBLEM DEFINTION:

Remote data integrity checking is of crucial importance in cloud storage. In multi-cloud environment, distributed provable data possession is an important element to secure the remote data. we propose a novel remote data integrity checking model: ID-DPDP (identity-based distributed provable data possession) in multi-cloud storage. The proposed ID-DPDP protocol is provably secure under the hardness assumption of the standard CDH (computational Diffi Hellman) problem. The proposed ID-DPDP protocol can realize private verification, delegated verification and public verification.

ADVANTAGES:

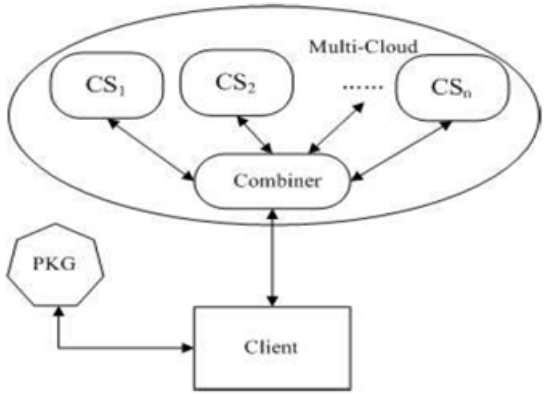
•The distributed cloud storage is indispensable.
•Efficient and Flexible.
•Elimination of the certificate management.

SYSTEM ARCHITECTURE:



ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The cloud servers respond the challenge and the combiner aggregates these responses from the cloud servers. The combiner sends the aggregated response to the verifier. Finally, the verifier checks whether the aggregated response is valid. The concrete ID-DPDP construction mainly comes from the signature, provable data possession and distributed computing. The signature relates the client's identity with his private key.



IMPLEMENTATION:

Multi cloud storages:

Distributed computing is used to refer to any large collaboration in which many individual personal computer owners allow some of their computer's processing time to be put at the service of a large problem. In our system the each cloud admin consist of data blocks .the cloud user upload the data into multi cloud. cloud computing environment is constructed based on open architectures and interfaces, it has the capability to incorporate multiple internal and/or external cloud services together to provide high interoperability. We call such a distributed cloud environment as a multi-Cloud .A multi-cloud allows clients to easily access his/her resources remotely through interfaces.

Cooperative PDP:

Cooperative PDP (CPDP) schemes adopting zero-knowledge property and three-layered index hierarchy, respectively. In particular efficient method for selecting the optimal number of sectors in each block to minimize the computation costs of clients and storage service providers. cooperative PDP (CPDP) scheme without compromising data privacy based on modern cryptographic techniques.
Data Integrity:-

Data Integrity is very important in database operations in particular and Data warehousing and Business intelligence in general. Because Data Integrity ensured that data is of high quality, correct, consistent and accessible.

PKG (Private Key Generator):

an entity, when receiving the identity, it outputs the corresponding private key.

Third Party Auditor:

Trusted Third Party (TTP) who is trusted to store verification parameters and offer public query services for these parameters. In our system the Trusted Third Party, view the user data blocks and uploaded to the distributed cloud. In distributed cloud environment each cloud has user data blocks. If any modification tried by cloud owner a alert is send to the Trusted Third Party.

Cloud User:

The Cloud User who has a large amount of data to be stored in multiple clouds and have the permissions to access and manipulate stored data. The User's Data is converted into data blocks. The data block is uploaded to the cloud. The TPA views the data blocks and Uploaded in multi cloud. The user can update the uploaded data. If the user wants to download their files, the data's in multi cloud is integrated and downloaded.

CONCLUSION:

We presented the construction of an efficient PDP scheme for distributed cloud storage. Based on Homomorphism verifiable response and hash index hierarchy, we have proposed a cooperative PDP scheme to support dynamic scalability on multiple storage servers. We also showed that our scheme provided all security properties required by zero-knowledge interactive proof system, so that it can resist various attacks even if it is deployed as a public audit service in clouds.Furthermore, we optimized the probabilistic query and periodic verification to improve the audit performance. Our experiments clearly demonstrated that our approaches only introduce a small amount of computation and communication overheads. Therefore, our solution can be treated as a new candidate for data integrity verification in outsourcing data storage systems.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

As part of future work, we would extend our work to explore more effective CPDP constructions. Finally, it is still a challenging problem for the generation of tags with the length irrelevant to the size of data blocks. We would explore such a issue to provide the support of variable-length block verification.

REFERENCES:

[1] Huaqun Wang, Identity-Based Distributed Provable Data Possession in Multicloud Storage, IEEE Transactions on Services Computing, (Volume:8 , Issue: 2)

[2] G. Ateniese, R. DiPietro, L. V. Mancini, G. Tsudik, "Scalable and Efficient Provable Data Possession", SecureComm 2008, 2008.

[3] C. C. Erway, A. Kupcu, C. Papamanthou, R. Tamassia, "Dynamic Provable Data Possession", CCS'09, pp. 213-222, 2009.

[4] F. Seb´e, J. Domingo-Ferrer, A. Mart´ınez-Ballest´e, Y. Deswarte, J.Quisquater, "Efficient Remote Data Integrity checking in Critical Information Infrastructures", IEEE Transactions on Knowledge and Data Engineering, 20(8), pp. 1-6, 2008.

[5] H.Q. Wang, "Proxy Provable Data Possession in Public Clouds," IEEE Transactions on Services Computing, 2012. http://doi.ieeecomputersociety.org/10.1109/TSC.2012.35

[6] Y. Zhu, H. Hu, G.J. Ahn, M. Yu, "Cooperative Provable Data Possession for Integrity Verification in Multicloud Storage", IEEE Transactions on Parallel and Distributed Systems, 23(12), pp. 2231-2244, 2012.

[7] Y. Zhu, H. Wang, Z. Hu, G. J. Ahn, H. Hu, S. S. Yau, "Efficient Provable Data Possession for Hybrid Clouds", CCS'10, pp. 756-758, 2010.

[8] R. Curtmola, O. Khan, R. Burns, G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession", ICDCS'08, pp. 411-420, 2008.

[9] A. F. Barsoum, M. A. Hasan, "Provable Possession and Replication of Data over Cloud Servers", CACR, University of Waterloo, Report2010/32,2010. Available at http://www.cacr.math.uwaterloo.ca/techreports /2010/cacr2010-32.pdf.

[10] Z. Hao, N. Yu, "A Multiple-Replica Remote Data Possession Checking Protocol with Public Verifiability", 2010 Second International Symposium on Data, Privacy, and E-Commerce, pp. 84-89, 2010.

[11] A. F. Barsoum, M. A. Hasan, "On Verifying Dynamic Multiple Data Copies over Cloud Servers", IACR eprint report 447, 2011. Available at http://eprint.iacr.org/2011/447.pdf.

[12] A. Juels, B. S. Kaliski Jr., "PORs: Proofs of Retrievability for Large Files", CCS'07, pp. 584-597, 2007.

[13] H. Shacham, B. Waters, "Compact Proofs of Retrievability", ASIACRYPT 2008, LNCS 5350, pp. 90-107, 2008.

[14] K. D. Bowers, A. Juels, A. Oprea, "Proofs of Retrievability: Theory and Implementation", CCSW'09, pp. 43-54, 2009.

[15] Q. Zheng, S. Xu. Fair and Dynamic Proofs of Retrievability. CODASPY' 11, pp. 237-248, 2011.

[16] Y. Dodis, S. Vadhan, D. Wichs, "Proofs of Retrievability via Hardness Amplification", TCC 2009, LNCS 5444, pp. 109-127, 2009.

[17] Y. Zhu, H. Wang, Z. Hu, G. J. Ahn, H. Hu, "Zero-Knowledge Proofs of Retrievability", Sci China Inf Sci, 54(8), pp. 1608-1617, 2011.

[18] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing", INFOCOM 2010, IEEE, March 2010.

[19] Q. Wang, C. Wang, K. Ren, W. Lou, J. Li, "Enabling Public Audit ability and Data Dynamics for Storage Security in Cloud Computing", IEEE Transactions on Parallel And Distributed Systems , 22(5), pp. 847-859,2011.

[20] C. Wang, Q. Wang, K. Ren, N. Cao, W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing," IEEE Transactions on Services Computing, 5(2), pp. 220-232, 2012.