ISSN No. 0976-5697

Volume 4, No. 4, March-April 2013

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Framework of Data Integrity for Cross Cloud Environment Using CPDP Scheme

Sarita Motghare*
G.H.R.I.E.T.W, RTMNU Nagpur
saritamotghare@yahoo.com

P.S.Mohod
G.H.R.I.E.T.W, Nagpur
psmohod@gmail.com

S.P.Khandait
K.D.K.C.E, Nagpur
prapti_khandait@yahoo.co.in

Anil Jaiswal
G.H.R.I.E.T.W, Nagpur
jaiswal.anil@gmail.com

Abstract: In recent years, cloud storage service has become a faster profit growth point by providing a comparably scalable, position-independent, low-cost platform for client's data. Since cloud computing environment is constructed based on open architectures and interfaces. It has the capability to incorporate multiple internal and external cloud services together to provide high interoperability there can be multiple accounts associated with a single or multiple service providers (SPs).so, Security in terms of integrity is most important aspect in cloud computing environment.

Cooperative Provable data possession (CPDP) is a technique for ensuring the integrity of data in storage outsourcing. Therefore, we address the construction of an efficient CPDP scheme and dynamic audit service for distributed cloud storage as well verifying the integrity guarantee of an entrusted and outsourced storage which support the scalability of service and data migration.

Keywords**:** Cloud computing, Cooperative Provable data possession, Data storage, Integrity verification, Proof of retrievability.

I. INTRODUCTION

Many trends are opening up the era of Cloud Computing, which is an Internet-based development and use of computer technology. Consider the large size of the outsourced electronic data and the client's constrained resource capability, the core of the problem can be generalized as how can the client find an efficient way to perform periodical integrity verifications without the local copy of data files.

The main objective of this paper is to provide security in terms of integrity and availability of client's data which is stored on cloud. This paper shall not put any burden on to computation and communication and further, performance guarantee shall also be taken care of by allowing trusted third party to verify the correctness of the cloud data on demand without retrieving a copy of the whole data or introducing additional on-line burden to cloud users.

Several schemes [1][2][7] are proposed to solve the problem. Those schemes focus on achieve the following requirements: high efficiency, stateless verification, retrievability of data, unbounded use of queries and public verification. In general, if one scheme supports private verification, it can possess higher efficiency.

II. MOTIVATION

Because the speed of today's data has produced far more than the current availability of storage devices, so there will be more and more data need to be outsource [2].The cloud computing has been seen as the next generation of enterprise IT infrastructure, software, applications as a service and users will also concentrated all the information stored in the cloud data centre, this new data storage model will bring new challenges and new problems. One of the most important and most attention issues, that is in the cloud environment, servers within the data storage with security in terms of integrity verification. For example, storage service providers may order their own interests to save the data to hide an error, more seriously, storage service providers in order to save cost and storage space, deliberately remove rarely accessed data, and then who, due to extensive confidential information, outsourcing and limited computing power users. Therefore, how to backup data files in the user not the case, found an efficient and securely ways of good information to perform periodically verification, allowing users to know his information file is stored securely on the server, this data storage is cloud computing environment is an important security issue.

III. CONTRIBUTION

Our proposed agreement has two main contributions

a. ***Efficiency and Security:*** the plan proposed by the CPDP [1][2] is safer to rely on a public and private key encryption will be clear, efficient in the use of SecretKeyGen and TagGen[5] algorithms. In this every time parameters are generated and key exchange takes place so more secure than symmetric and asymmetric algo. However, our plan is more efficient than the other techniques. Because it does not require lots of data encryption in outsourced and no additional posts on the symbol block, and the ratio [8] is more secure because we encrypt data to prevent unauthorized third parties to know its contents.

b. ***Public verifiability:*** We plan a major variation of CPDP, to provide public validation. Allow people other than the owner for information on the server has proved challenge. However, our program than [2] is more efficient because it does not need the information for each block encryption.

Paper structure Framework for the rest of the paper is as follows. In section IV, we describe the related work. Section V describes a data integrity for cross cloud environment using CPDP scheme to prove a structure, emphasizing the characteristics of CPDP and the related parameters. In section VI, we introduce the CPDP can be publicly verifiable information to prove a structure.Then, Section VII security analysis of our protocol, and VIII is about results and graphs However, Section IX is our conclusion.

IV. RELATED WORK

Nam Yem Li et al. [2], highlights PDP scheme use for verification to avoid public verification. This paper proposed initial PDP solution to RSA based Hash function to authenticate the remote server storage data. However, due to RSA based cryptosystem, the entire computing speed is slow.

Similarly Qian Wang et al. [7], Proposes a protocol for Integrity verification in Multi cloud that is provided by improving the existing proof of storage models by manipulating the classic Merkle Hash Tree construction for block tag authentication. To support efficient handling of multiple auditing tasks, this paper further explore the technique of bilinear aggregate signature to extend the main result into a multiuser setting, where TPA can perform multiple auditing tasks simultaneously. Extensive security and performance analysis show that the proposed schemes are highly efficient and provably secure. This paper explored the problem of providing simultaneous public audibility and data dynamics for remote data integrity check in Cloud Computing. This Study mproves the existing proof of storage models by manipulating the classic Merkle Hash Tree construction for block tag authentication. major concern of this paper is, It is used to construct verification protocols that can accommodate dynamic data files.

Then, Yan Zhu et al. [3], gives Collaborative Provable Data Possession scheme, where collaborative integrity verification mechanism in hybrid clouds to support the scalable service and data migration, in which we consider the existence of multiple cloud service providers to cooperatively store and maintain the client's data. This paper is for a hybrid cloud is a cloud computing environment in which an organization provides and manages some internal resources and the others provided externally. The performance optimization mechanisms scheme is satisfactory and proves the security of that scheme based on multi-proven zero-knowledge proof system, which can satisfy the properties of completeness, knowledge soundness, and zero-knowledge.

Subsequently Yan Zhu et al. [1], focuses on the Cooperative Provable data possession scheme for integrity verification. This scheme is based on homomorphic verifiable response and hash index hierarchy for data access. This paper issued, to prove the Security of scheme based on multi-prover zero knowledge proof system. CPDP scheme provides Integrity with lower computation and communication overheads in comparison to non cooperative approach. However, while checking for large files, integrity is affected by the bilinear mapping operations due to its high complexity. And generation of tags with the length irrelevant to the size of data blocks is a challenging task of this paper.

In the literature [1], proposed a data storage proved cooperative Provable Data Possession (CPDP) system, which applies to of cloud in an entrusted storage server, based on Diffie-Hellman protocol systems of main plant with state verify that the label is used to check the integrity of the data stored in the cloud, which allows unlimited number of storage server authentication, and also provides a public authentication method, In which the use of public and private key system and the data must be calculated when private key matches and tags the action, making it a relatively large amount of computation. Compared to the literature [1] of CPDP protocol, the literature [3] for the previous method proposed by CPDP [1] extension of a new dynamic storage technology, because, in this new method uses the Diffie-Hellman cryptography to encrypt, making information storage, bandwidth and computational smaller, more efficient. However, we found that in the actual case, verify the number is not a difficult problem. Therefore, our protocol is based on hybrid cryptography, so our protocol than the literature [1] more efficient than the literature [3] and more security, but also increase the public verification function.

The protocol is similar to the CPDP, Yan Zhu [1] proposed a Proof of retrievability [1] (PORs) system, and thus the system made many accurate proof and verification, in this system, the sampling code and error correction codes are also used to confirm the data on the control and verification, which more special place, purposes is to detect and block some random recessed special information block, and in order to protect those special blocks position, further use of asymmetric encryption technology. Compared to PORs [1], we proposed protocol requires less data storage space and use less bandwidth.

V. AN INTEGRITY FOR CROSS CLOUD ENVIRONMENT USING CPDP SCHEME

In this Agreement, the password system based on CPDP , the main idea is to outsource the file before the data block encryption, and validation of fixed-size tags, each tag are included in the block information. Fig. 1 is a Cross cloud environment in CPDP agreement setting the stage diagram:

Although existing CPDP schemes [1] offer a publicly accessible remote interface for checking and managing the tremendous amount of data, the majority of existing CPDP schemes are incapable to satisfy the inherent requirements from multiple clouds in terms of communication and computation costs. To address this problem, we consider a multi-cloud storage service as illustrated in Figure 1.

A. System Architecture:

In this architecture, a data storage service involves three different entities: Clients who have a large amount of data to be stored in multiple clouds and have the permissions to access and manipulate stored data Cloud Service Providers (CSPs) who work together to provide data storage services and have enough storages and computation resources. and Trusted Third Party (TTP) who is trusted to store verification parameters and offer public query services for these parameters.

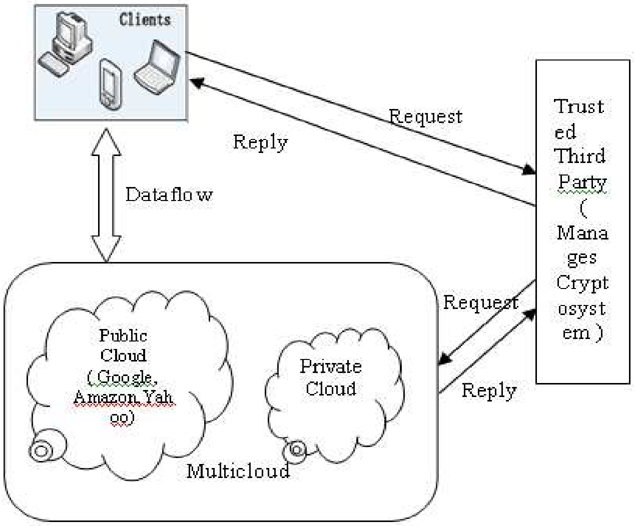

Clients
Request
Reply
Dataflow
Trusted Third Party (Manages Cryptosystem)
Request
Reply
Public Cloud (Google, Amazon,Yahoo)
Private Cloud
Multicloud

Figure 1: Verification of integrity in cross cloud environment

In this architecture, we consider the existence of multiple CSPs to cooperatively store and maintain the clients' data. Moreover, a cooperative PDP is used to verify the integrity and availability of their stored data in all CSPs. The verification procedure is described as follows: Firstly, a client (data owner) uses the secret key to pre-process a file which consists of a collection of n blocks, generates a set of public verification information that is stored in TTP, transmits the file and some verification tags to CSPs, and may delete its local copy; Then, by using a verification protocol, the clients can issue a challenge for one CSP to check the integrity and availability of outsourced data with respect to public information stored in TTP.

B. *Protocol Directions:*

We neither assume that CSP is trust to guarantee the security of the stored data, nor assume that data owner has the ability to collect the evidence of the CSP's fault after errors have been found. To achieve this goal, a TTP server is constructed as a core trust base on the cloud for the sake of security. We assume the TTP is reliable and independent through the following functions: to setup and

a. SecretKeyGen **(**1^k**):** Takes a security parameter k as input, and returns a secret key Sk or a public-secret keypair (Pk, Sk);

b. VeriTagGen **(**Sk,F,P**):** Takes as inputs a secret key Sk, a file F, and a set of cloud storage providers P = {Pk}, and returns the triples (S_t,V_p,A_t), where S_t is the secret in tags, V_p = (u,H) is a set of verification parameters u and an index hierarchy H for F, A_t = { A_t $^{(k)}$ bn} Pk belongs to P) denotes a set of all tags, A_t maintain the CPDP cryptosystem; to generate and store data owner's public key; and to store the public parameters used to execute the verification protocol in the CPDP the fraction $F^{(k)}$ of F in Pk.$^{(k)}$ is the verification tag of scheme. Note that the TTP is not directly involved in the CPDP scheme in order to reduce the complexity of cryptosystem [1].

(Cooperative-PDP).[4] A Cooperative provable data possession scheme S' is a collection of two algorithms And an interactive proof system, S' = (K, T, P):

The verification procedure is described as follows: Firstly, a client (data owner) uses the secret key (generated by proposed algorithm SecretKeyGen) to pre-process a file which consists of a collection of n blocks, generates a set of public verification information (generated by proposed VeriTagGen algorithm) that is stored in TTP, transmits the file and some verification tags to CSPs, and may delete its local copy. Then, by using this verification protocol, the clients can issue a challenge for one CSP to check the integrity and availability of outsourced data with respect to public information stored in TTP.

c. In proposed system a cooperative provable data possession in cross cloud S= (SecretKeyGen, VeriTagGen, proof) is a collection of two algorithms (SecretKeyGen, VeriTagGen) and an interactive proof system proof, as follows:
Proof**(**P,V**):** Is a protocol of proof of data possession between CSPs (P = {Pk) and a verifier (V), that is, < Σ Pk E P $P^{(k)}$, $F^{(k)}$,V_p $^{(k)}$ (Pk, V_{p},) where Pk takes input file $F^{(k)}$ and a set of tags u(k), and a public key pk and a set of public parameters V_p is the common input between P and V. At the end of the protocol run, V returns a bit {O / 1} denoting false and true.

Proposed work neither assumes that CSP is trust to guarantee the secusrity of the stored data, nor assume that data owner has the ability to collect the evidence of the CSP's fault after errors have been found. To achieve this goal, a TTP server is constructed as a core trust base on the cloud for the sake of security. Proposed work assume the TTP is reliable and independent through the following functions [1]: to setup and maintain the CPDP cryptosystem; to generate and store data owner's public key; and to store the public parameters used to execute the verification protocol in the CPDP scheme. But the TTP is not directly involved in the CPDP scheme in order to reduce the complexity of cryptosystem. This is proposed cross cloud scheme for key generation, tag generation and verification protocol..

C. *flow chart of proposed system:*

This proposed system on client side will work for two conditions for storing data request (SDR) and for Accessing data request (ADR). If client want to store data, with the help of TTP

Secret key is generated, by using that secret key data get stored. For accessing data, First TTP check for trust between clouds and then check for trusted kye between client and TTP, And user will get data. As shown in fig 2, the flow chart of the proposed system.

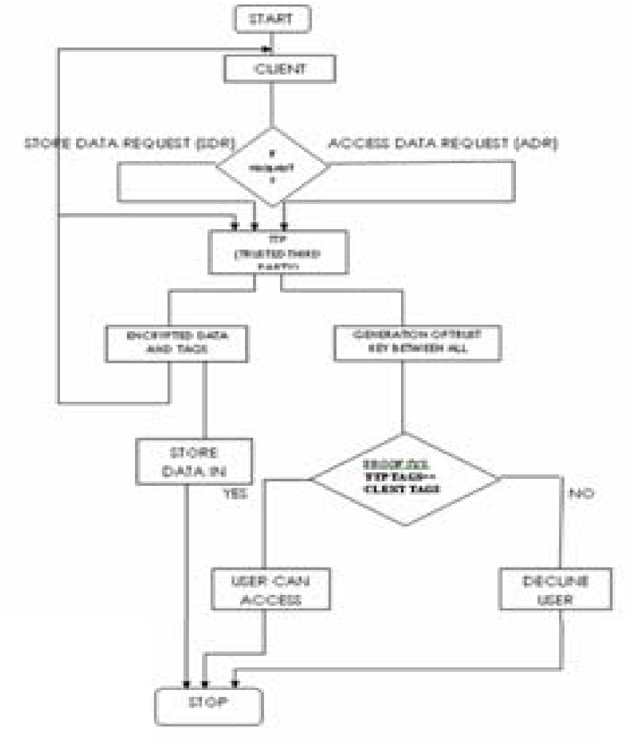


Figure 2: Flow chart for proposed system

VI. CPDP CAN BE PUBLICLY VERIFIABLE INFORMATION TO PROVE A STRUCTURE

If you use public verification, the owner (client) can be verified (challenge) phase separation, will want to verification task of entrusted to a third party to perform, but also because the owner of a trusted third party may have to better and more efficient than the hard physical equipment and computing power, it can improve the efficiency of verification. At this point, the owner do not have to produce itself verification of the server, do not have to verification proof of the value coming from the server, only to appoint tasks to a third party, which greatly reduce the owner's cost of computing and storage costs. Therefore, we further modify the static PDP protocol, and this stage is a hybrid variation of the static type of PDP is to provide publicly verify the characteristics of this phase can allow anyone to verify the correctness of data stored on the server.

A trusted third party can be calculated for each round of exchange of keys i k and the current challenges i c and to calculate the t times may be random challenges can be made to the server verification requirements. The publicly verify the hybrid static PDP hybrid verification phase of and the verification phase of the same static PDP, so we will not go into detail here.

Publicly available through our proposed cooperative PDP verification mechanism that allows information to authorized third parties for possession verification. However, due to the data file is encrypted by the data owner stored on the server in the cloud, so the data owners need to worry about his information in the authorized trusted third-party validation data was stolen or know the contents.

VII. SECURITY ANALYSIS

This section will analyze the static PDP hybrid security agreement to confidentiality, integrity and confirm the analysis of three aspects.

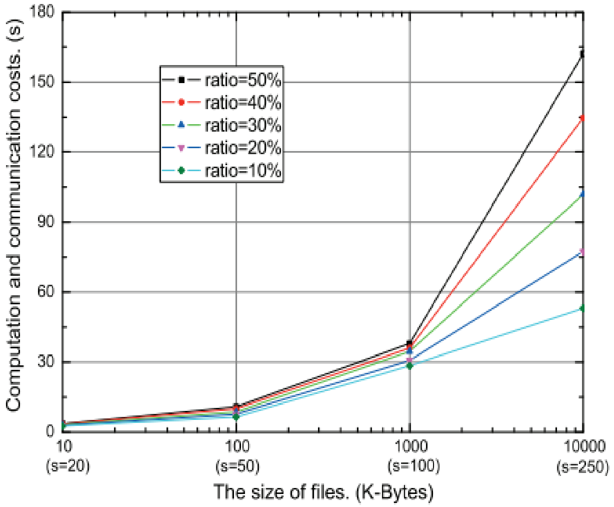
A. Confidentiality:

The owner of the file is stored on the server before, will use the cryptosystem to encrypt the data to ensure that the file will not be intercepted by an unauthorized person to get the file content. Because encryption and decryption *SecretKeyGen and VeriTagGen* cryptosystem uses public key and private key, security is based on calculating private key, Until and unless you don't know private key, you can't decrypt the ciphertext file M.

B. Integrity:

In the verification phase, the owner would like to verification ciphertext M is a complete file stored on the server at this time, the server will calculate the value of z to prove he has complete store ciphertext file M . If the server is calculated z calculated with the owner of the verification value is equal to V , it means the server does have the correct storage ciphertext file M .

VIII. RESULTS AND GRAPHS



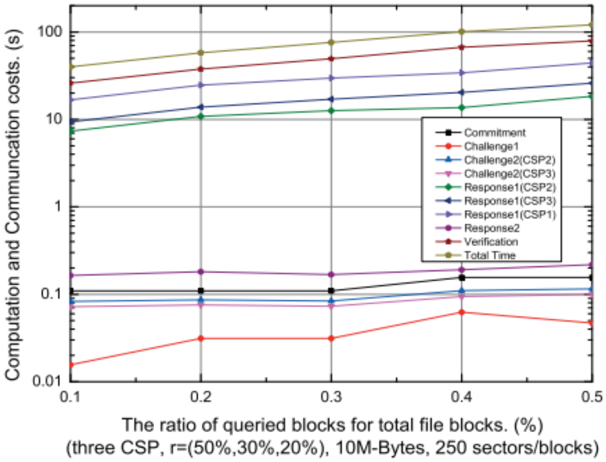


Figure 3 a,b)Experimentation result under different file size & Sampling ratio

As part of future work, we would extend our work to explore more effective CPDP constructions. First, from our experiments we found that the per-formance of CPDP scheme, especially for large files, is affected by the bilinear mapping operations due to its high complexity. To solve this problem, RSA-based constructions may be a better choice, but this is still a challenging task because the existing RSA-based schemes have too many restrictions on the performance and security [1]. Next, from a practical point of view, we still need to address some issues about integrating our CPDP scheme smoothly with existing systems, for example, how to match index-hash hierarchy with HDFS's two-layer name space, how to match index structure with cluster-network model, and how to dynamically update the CPDP parameters according to HDFS' specific requirements. Finally, it is still a challenging problem for the gener-ation 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.

IX. CONCLUSION

We focused the core issues, if an untrusted server to store customer information. We can use cooperative provable data possession scheme, which reduce the data block access, and amount of computation on the server and client. Also decreases server traffic.

Our design and development on the CPDP program is mainly based on the usage of Public and Private keyencryption system. It exceeds what we did in the past, the improvement has brought to the bandwidth, computation and storage system. And it applied the public (trusted third party) verification.\ Finally, we also expect our program, it supports dynamic outsourcing of information make it a more realistic application of cloud computing environment.

X. REFERRENCES

[1]. Yan Zhu, Hongxin Hu, Gail-Joon Ahn,"Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage" IEEE Transactions On Parallel And Distributed Systems,Digital Object Indentifier 10.1109/TPDS 2012.66 April 2012.

[2]. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner, Z. N. J. Peterson, and D. X. Song, "Hybrid Provable Data Possession at Untrusted Stores In Cloud Computing," in IEEE Conference on the 7th International Conference On Parallel And Distributed Systems 10.1109/ICTPDS 2011.70.

[3]. Y. Zhu, H. Hu, G.-J. Ahn, Y. Han, and S. Chen,"Collaborative integrity verification in hybrid clouds," in IEEE Conference on the 7th International Conference on Collaborative Computing: Networking Applications and Worksharing, collaborateCom, Orlando,Florida, USA, October 15-18, 2011, pp. 197–206

[4]. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S. Yau, "Study on the Third-party Audit in Cloud Storage Service s," in IEEE TRANSACTIONS ON SERVICES COMPUTING, Digital Object Indentifier 10. 1109/ TCS.2011.51.

[5]. Qian Wang,, Cong Wang, Kui Ren, Wenjing Lou, and Jin Li "Dynamic audit services for integrity verification of outsourced storages in clouds", VOL. 22, NO. 5, MAY 201, 1045- 9219/11/$26.00 2011 IEEE. 10.1109/ IMCCC. 2011.135.

[6]. Qian Wang, Cong Wang, Kui Ren, Wenjing Lou, and Jin Li "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing" IEEE Transactions on Parallel And Distributed Systems, VOL. 22, NO. 5, MAY 2011.