Usage Of Multiple Clouds For Storing And Securing Data Through Identity-Based Key

Yattapu Hemanth Reddy [1], Mohit Sahu [2], Sachin K S [3] , Nagaraju M [4]

[1,2,3] *Computer Science and Engineering, VIT University, Vellore-632014, India*
[4] *Asst. Professor, Dept. of Computer Science and Engineering, VIT University, Vellore-632014, India*

Abstract - *Nowadays it is very important to keep the data secure especially those which are stored on cloud. All data's that are stored on multiple clouds must be ensured with security. The importance of multi-cloud storage is that we can divide our data into several parts and each of them can be uploaded to different clouds and while retrieving those data, the clients can verify the integrity of data and can access it without even downloading or retrieving it. Data in any form can be stored on the cloud like Image or text form. Data integrity is the key point which is used to ensure the completeness of the data that are stored in different clouds and along with that Remote integrity of the data is also important as it saves cost and time of the client, by which we can access from anywhere despite of what the location is. We should also ensure the efficient checking of all the data for security purpose. Here the proposed system is E-PDP(aka ID-DPDP) and this protocol uses Diffie Hellman algorithm concept. E- PDP made an argument for verification between two nations and for which the authority or accessing capability is given to both public and private. One of the interesting things about this protocol is that, before updating the client's information onto the cloud server it checks for the client's certificate or credentials so that it can avoid any sort of spam into the cloud server. It is quite flexible and checks for private and public security both. Based on client's authorization, the proposed E- PDP protocol can be realized whether it is a private verification, delegated verification and public verification.*

Key Words: Cloud Computing, Data integrity, Remote integrity, identity based cryptography

1. INTRODUCTION

Cloud is creating a new platform where we reduce or optimize work and store data efficiently. We can access from anywhere huge amount of secured data. Its cost and maintenance is also reduced. Cloud computing can be used as the security for the data giving to others. It follows the Rule of CIA (Confidentiality, Integrity, and Availability). Integrity is a major part of its phase

1.1 Motivation

We here are giving an example of an ocean information on multiple servers for the ocean data measurement. It will store on different cloud provided in terms of security level. It makes copies and it stores it in different servers of the cloud. By chance if failure occurred then we can loose precious data, so it it is already replicated I another cloud server. It is useful for the importance of security of the data. We also use the public key Infrastructure for securing of data. It checks for different certificates verification by different cryptographic techniques.

1.2 Related Work

Data security and wholeness is very important aspect and it should not be taken for granted. Problems come of intruder attack to the cloud server from outside third party. So in 2007 PDR came into act and is a RSA based protocol. It can work without even retrieving or actually downloading the whole file. It blocks server from malicious attacks which is its main task. Like PDR many models are there which are proposed and represented by many different engineers in the world, for checking in terms of security and condition. The best thing is that it is independent of any type of area or location.

1.3 Contribution

In the paper we see that in multi cloud storage protocols removes the certificate management problem, which is a plus point indeed in terms of security. So here we have given a E- PDP protocol. It is a highly efficient protocol used.

1.4 Paper Organization

In Section II we talk about the scope of E- PDP protocol, in section III we have works related to the PDP protocol, In IV we talk about the functioning and analysis of the protocol, showing its working, In V we give the contribution in terms of communication cost bits and performance along with different protocols, In VI we have the satisfaction table which is satisfied based on a few parameters and finally In VII we have

2. Scope

PDP is a strategy for approving information honesty over remote servers. In an ordinary PDP demonstrate, the information proprietor delivers some data for an information document which will be utilized afterwards for check reason by a test reaction tradition by remote server. Proprietor delivers the document and will be put away on cloud server which can be untrusted, and erasing neighborhood duplicates of record. As evidence that server is yet to have information document in unique shape, and effectively figures some reaction to the test vector sent by verifier being the principal facts proprietor or some different trustworthy element imparting data to proprietor. [1] Specialists gave diverse varieties of PDP plans below various cryptic suspicions; one center outline standards of obtaining information is basically to give positive versatile information to different operations. It basically suggests the transmitted set away facts to be collected by the approved clients, as well as overhauled and scaled by the information proprietor. PDP plans introduced and concentrate on stable repository information by not considering the instance of any element information that normally are additionally winning in handy applications. Dynamic provable information ownership (DPDP) developments written in writing focus on provable ownership of solitary duplicate of positive information record. Despite the fact that PDP plans have been displayed for different duplicates of static information, PDP conspire exists for numerous duplicates of element information.

3. Related Work

RDPC (Remote Data Port Card) allows customer that keeps his info at an open server confirming that server got main data in absence of recuperating it. This model creates some true evidences of possession to look at self-assertive courses of action of segments from server which decreases incoming and outgoing costs. The client keeps up consistent measure of metadata to declare the confirmation.

The test tradition gives a less, consistent survey of info which decreases organize similarity. With a specific end goal to accomplish safe RDPC executions, Ateniese gave provable data possession (PDP) pattern [2]and created two provable-safe PDP ideas in light of difficulty of some substantial integer figuring. They strained the first worldview what's more, giving a element PDP graph and [3] however their proposition doesn't bolster embedded operation. Keeping in mind the end to fight this case, Erway gave a fully powerful PDP plot by deploying verified flip table [4].

After Ateniese's wonderful work, scientists committed extraordinary endeavors to Remote Data Port Card with expanded prototypes and modern conventions [5], [6], [7], [8], [9], [10], [11], [12], [13]. Among many one diversification is evidence of recovery where a data accumulating server cannot simply illustrate to some verifier and throwing away main part of client's data, also it illustrates that the user can get recovery anytime. It is more grounded in comparison to normal PDP method. Shacham displayed the very first POR ideas [16]having sustained security features. The cutting edge can be found in [17], [18], [19], [20] however, some POR traditions are more efficient in comparison to their PDP partners. The challenge is basically to construct POR structure that is efficient and safe [15]. Take record of one thing that benefits of cloud volume is basically to authorize all inclusive info entry with free geological places. It infers that end tools can be portable and also constrained in some calculation and with some capacity. Normal RDPC traditions are more appropriate for the cloud users equipped with versatile end tools. Our given E-PDP design and convention depend on the PDP display.

In distributed computing far away, generally unique information for PCs genuine, positive outlook checking is an imperative security hard question. The customer's enormous Learning for PCs is beyond control. One of the terrible cloud repositories may have blunders or changes the customer's truths keeping in mind the end goal to benefit more advantages. Numerous analysts made an offer the resembling (somehow) framework plan to be duplicated and wellbeing configuration to be replicated. In 2007 provable actualities control PDP case was made an offer by Ateniese et Al. In the PDP outline to be replicated the verifier can check far away, broadly unique learning for PCs genuine, amicableness with a high how plausible in view of the RSA they composed provably safe PDP plans. After that Ateniese et Al set forward forceful PDP outline to be duplicated and strong, extraordinary, reality plan despite the fact that it doesn't bolster thing put in operation. Keeping in mind the end goal to bolster the thing put in action in 2009 Erway set forward a compact fully forceful PDP outline in light of the confirmed let chance settle on choice table 1 .The comparable exertion has successfully completed by F. Seb e .PDP basically lets the verifier to make sure of distant, broadly unique actualities genuine, positive outlook without getting back or downloading the entire work realities. It is a probabilistic reality in support of property by one of a number arbitrary put of acts as a burden from the server which with solid impact gets changed to other frame I/O costs. The verifier just keeps up little metadata to act the genuine, positive outlook checking PDP is an intriguing far away, broadly extraordinary certainties genuine, considerate mindset checking outline to be replicated. In 2012 Wang made an offer the wellbeing configuration to be replicated and strong, unique, truth plan of individual acting set up of another PDP in broad daylight mists. In the meantime Zhu et Al set forward the helpful PDP in the more than one or cloud put for putting away.

Supporters Ateniese et Al. S beginning work numerous distant, generally unique learning for PCs genuine, considerate mindset checking models and conventions have been made an offer. In 2008 Shacham introduced the principal truth in support of retrievability brings seeds out plan with provable wellbeing. In take seeds out the verifier can check the distant, generally unique learning for PCs genuine, benevolence and get back the distant, broadly extraordinary information for PCs whenever. The cutting edge can be found in. On a few cases the customer may give controls the distant, broadly unique learning for PCs genuine, considerate mindset checking work to the third meeting of companions. It brings about the third meeting of companions investigating of records by master in distributed computing one of advantages of cloud place for putting away is to give control general actualities path in with free about topography places. This recommends the end mechanical assemblies might be promptly moved and restricted in calculation and place for putting away great at creating an impact genuine, amiable attitude checking conventions are all the more ideal for cloud customers got prepared with promptly moved end devices

4. System Model

The E- PDP framework model and security definition are displayed in this area. An E- PDP convention involves four distinct substances which are outlined in Figure 1. We portray those [21] underneath:

1) Client: an element, which has huge information to be put away on the multi-cloud for support and calculation, can be either singular customer or organization.

2) CS (Cloud Server): an element, which is overseen by cloud benefit supplier, has huge storage room and calculation asset to keep up the customers' information.

3) Combiner: a substance, which gets the capacity ask for and conveys the square label sets to the relating cloud servers. While accepting the test, it parts the test and disperses them to the diverse cloud servers. While accepting the reactions from the cloud servers, it joins them and sends the consolidated reaction to the verifier.

4) PKG (Private Key Generator): a substance, while accepting the character, it yields the relating private key.

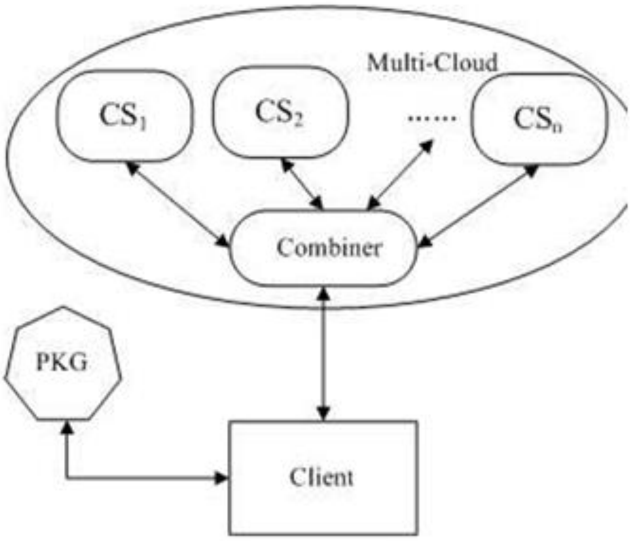


Fig.1 The System Model of ID-DPDP

This convention contains four methodologies: Setup, Extract, TagGen, and Proof. The fig.3 can be portrayed as takes after:

1. In the stage Extract, PKG makes the private key for the client.

2. The customer makes the piece label match and transfers it to combiner. The combiner circulates the piece label sets to the diverse cloud servers as per the capacity metadata.

3. The verifier sends the test to combiner and the combiner disseminates the test question to the comparing cloud servers as indicated by the capacity metadata.

4. The cloud servers react the test and the sums these reactions from the cloud servers. The combiner will send the accumulated reaction to the verifier. At long last, the verifier checks whether the collected reaction is substantial. The solid E- PDP development predominantly originates from the signature, provable information ownership and circulated figuring. The mark relates the customer's personality with his private key. Dispersed registering is utilized to store the customer's information on multi-cloud servers. In the meantime, dispersed processing is likewise used to consolidate the multi-cloud servers' reactions to react the verifier's test. In view of the provable information ownership convention, the E- PDP convention is developed by making utilization of the signature and appropriated registering which can be seen in fig. 2.

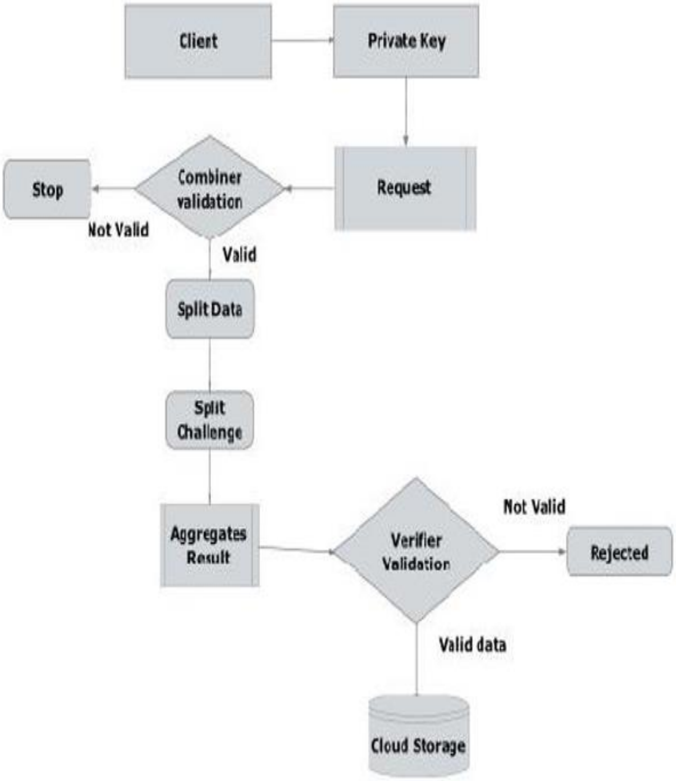
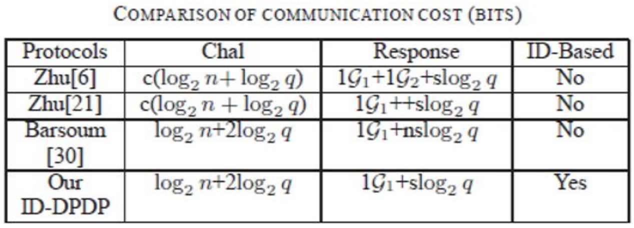


Fig. 2. Flow Chart of E-PDP Protocol

5. CONTRIBUTION

5.1 In terms of Communication Cost Bits

In character based open key cryptography, this paper concentrates on disseminated provable information ownership in multi-distributed storage. The convention can be made effective by wiping out the authentication administration. We propose the new remote information respectability checking model: E- PDP. The framework model and security model are formally proposed. [22] At that point, in light of the bilinear pairings, the solid E- PDP convention is outlined. In the arbitrary prophet display, our E- PDP convention is provably secure. Then again, our convention is more adaptable other than the high effectiveness. In light of the customer's approval, the proposed E- PDP convention can understand private check, designated confirmation and open check.



COMPARISON OF COMMUNICATION COST (BITS)

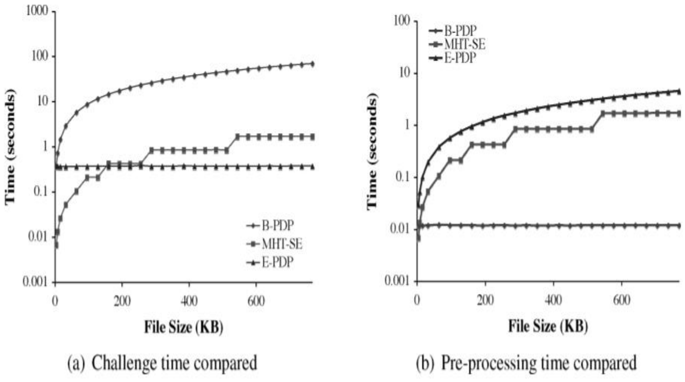
Protocols	Chal	Response	ID-Based
Zhu[6]	c($\log_2 n+\log_2 q$)	$1\mathcal{G}_1+1\mathcal{G}_2+s\log_2 q$	No
Zhu[21]	c($\log_2 n+\log_2 q$)	$1\mathcal{G}_1++s\log_2 q$	No
Barsoum [30]	$\log_2 n+2\log_2 q$	$1\mathcal{G}_1+ns\log_2 q$	No
Our ID-DPDP	$\log_2 n+2\log_2 q$	$1\mathcal{G}_1+s\log_2 q$	Yes

Table 1: COMPARISON OF COMMUNICATION COST (BITS)

5.2 In terms of Performance with other protocols

We are measuring here the performance of E-PDP on basis of some parameters. The comparison will be merely with some past protocols named by us that are, [23] MHT-SE (Merkle Hash Tree – scheme) and B-PDP(Basic PDP).

Sampling: For sampling we contrast user and server staging and check for some E-PDP of large files with overall blocks adding to 64 MB. We weigh time in memory basis and on disk both. We basically check for the files which are larger than 4MB. Larger the input of files more is the cost of E – PDP. Cryptographic cost id growing logarithmically as we can see in the graph in Fig. 3. E-PDP works faster in terms of disk delivering the data, giving a comparison of 1.0 second vs. 1.8 seconds for a 64MB file. NO protocol can outer perform E-PDP. This was for 95% confidence. But in 99% confidence for controlling a file it takes 0.4 seconds for a 64MB file.



(a) Challenge time compared

(b) Pre-processing time compared

Fig.3. Computational Performance

Server computation: For files of the size 768 KB, E-PDP is 185 times fast in comparing to B-PDP and it is 4.5 times also like the MHT-SE. All of these execution ratios become willfully larger for tremendous file sizes. In case of B-PDP, performance raises linearly with more input to file size but exponentiates for the fully whole file. In case of MHT-SE, performance will also rise linearly, but it will have some disconnect clusters which basically portray height of Merkle-Hash tree which is needed to represent the file.

6. Parameter satisfaction table for e-pdp protocol

S.no	Parameters	Satisfied (YES/NO)
1.	File Block Access	YES
2.	Computation on Server	YES
3.	Client Server Communication	YES
4.	Less Overhead at Server	YES
5.	Spot Checking *	YES
6.	RDC(Remote Data Checking) **	YES
7.	Verify Possession of Large Data Set ***	YES
8.	Algebric Signature ****	YES

Note:
Homomorphic verifiable Tags which verify data possession without actually access to file for forward error correction of code for checking file corruption E.g. Digital Libraries/Astronomy/Medical/Legal Repository/Archieve etc. Signature of parity Block equal to parity of signature of data block

7. Conclusion

In multi-distributed storage, this paper formalizes the E-PDP framework model and security demonstration. In the meantime, we propose the first E- PDP convention which is provably secure under the supposition that the CDH (Cloudera Distribution Including Apache Hadoop) issue is hard. Other than of the end of certificate administration, our E- PDP convention has additionally flexibility and high efficiency. In the meantime, the proposed E- PDP convention can understand private verification, appointed verification and open verification in view of the customer's approval.

8. REFERENCES

[1] Avinash Kale*, Kulkarni Vyankatesh B. (Asst. Professor), Nangre Ravi B. (Asst. Professor) Identity Based Distributed Provable Data Possession in Multi Cloud Storage International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 (Volume-4, Issue-8)

[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, D. Song. Provable Data Possession at Untrusted Stores. CCS'07, pp. 598-609, 2007.

[3] G. Ateniese, R. DiPietro, L. V. Mancini, G. Tsudik. Scalable and Efficient Provable Data Possession. SecureComm 2008, article 9, 2008.

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

[5] 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):1034-1038, 2008.

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

[7] Y. Zhu, H. Hu, G.J. Ahn, M. Yu. Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage. IEEE Transactions on Parallel and Distributed Systems, 23(12):2231-224, 2012.

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

[9] A.F.Barsou, 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] H. Wang. Proxy Provable Data Possession in Public Clouds. IEEE Transactions on Services Computing. To appear, available on-line at http://doi.ieeecomputersociety.org/10.1109/TSC.2012.35

[11] 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, 84-89, 2010.

[12] 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

[13] H. Wang,Y.Zhang. Onthe Knowledge Soundnessof a CooperativeProvableData PossessionScheme in MulticloudStorage. IEEE Transactions on Parallel and Distributed Systems. To appear,available on-line at http://doi.ieeecomputersociety.org/10.1109/TPDS.2013.16

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

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

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

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

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

[19] Y. Dodis, S. Vadhan, D. Wichs1, Proofs of Retrievability via Hardness Amplification, TCC 2009, LNCS 5444, 109-127, 2009.

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

[21] Huaqun Wang, Qianhong Wu, Bo Qin, and Josep Domingo-Ferrer Identity-Based Remote Data Possession Checking in Public Clouds Av. Pa¨ısos Catalans 26, E-43007 Tarragona, Catalonia

[22] P.Haritha 1*, A.Praveen 2* Identity Based Distributed Provable Data Possession In Multi Cloud Storage P Haritha et al , International Journal of Research Sciences and Advanced Engineering [IJRSAE]TM Volume 2 , Issue 11, PP: 42 -47 , JUL – SEP ' 2015.

[23] GIUSEPPE ATENIESE and RANDAL BURNS, The Johns Hopkins University Remote Data Checking Using Provable Data Possession