International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 8, August 2014)

Scalable & Efficient Provable Data Possession for Integrity Verification in Private & Hybrid Clouds

Nagaraj Peddarapu
Associate Professor
Sri Indu College of Engg & Technology
Ibrahimpatan, Hyderabad, Telangana, India
nagaraj.peddarapu@gmail.com

Dr. Ch G.V.N. Prasad
Professor & HOD of CSE Dept
Sri Indu College of Engg & Technology
Ibrahimpatan, Hyderabad, Telangana, India

Abstract: **Provable information possession (PDP) may be a technique for guaranteeing the integrity of knowledge in storage outsourcing. during this paper, we have a tendency to address the development of associate degree economical PDP theme for distributed cloud storage to support the measurability of service and information migration, within which we have a tendency to contemplate the existence of multiple cloud service suppliers to hand and glove store and maintain the clients' information. we have a tendency to gift a cooperative PDP (CPDP) theme supported homomorphic verifiable response and hash index hierarchy. we have a tendency to prove the protection of our theme supported multi-prover zero-knowledge proof system, which may satisfy completeness, data soundness, and zero- data properties. additionally, we have a tendency to articulate performance optimisation mechanisms for our theme, associate degreed specially gift an economical technique for choosing optimum parameter values to attenuate the computation prices of shoppers and storage service suppliers. Our experiments show that our resolution introduces lower computation and communication overheads compared with non-cooperative approaches.**

1. INTRODUCTION

In recent years, cloud storage service has become a quicker profit growth purpose by providing a comparably low-priced, scalable, position-independent platform for clients' information. Since cloud computing atmosphere is built supported open architectures and interfaces, it's the potential to include multiple internal and/or external cloud services along to produce high ability. we have a tendency to decision such a distributed cloud atmosphere as a multi-Cloud (or hybrid cloud). Often, by using virtual infrastructure management (VIM) , a multi-cloud permits shoppers to simply access his/her resources remotely through interfaces like net services provided by Amazon EC2. There exist numerous tools and technologies for multicloud, such as Platform VM arranger, VMware v Sphere, and Overt. These tools facilitate cloud providers construct a distributed cloud storage platform (DCSP) for managing clients' data. However, if such a vital platform is prone to security attacks, it would bring irrecoverable losses to the shoppers. for instance, the confidential information in Associate in Nursingenterprise could also be lawlessly accessed through a far off interface provided by a multicloud, or relevant information and archives could also be lost or tampered with once they area unit stored into Associate in Nursing unsure storage pool outside the enterprise. Therefore, it isindispensable for cloud service suppliers (CSPs) to produce security techniques for managing their storage services. Demonstrable information possession (PDP) [2] (or proofs of retrievability (POR)) is such a probabilistic proof technique for a storage supplier to prove the integrity and possession of clients' information while not downloading information. The proof-checking while not downloading makes it particularly necessary for large-size files and folders (typically as well as several clients' files) to see whether or not these information have been tampered with or deleted while not downloading the most recent version of data. Thus, it is able to exchange ancient hash and signature functions in storage outsourcing. Various PDP schemes are recently projected, like climbable PDP and Dynamic PDP. However, these schemes primarily specialise in PDP problems at un trusted servers in an exceedingly single cloud storage supplier and aren't appropriate for a multi-cloud environment.

1.1. Verification architecture for data integrity

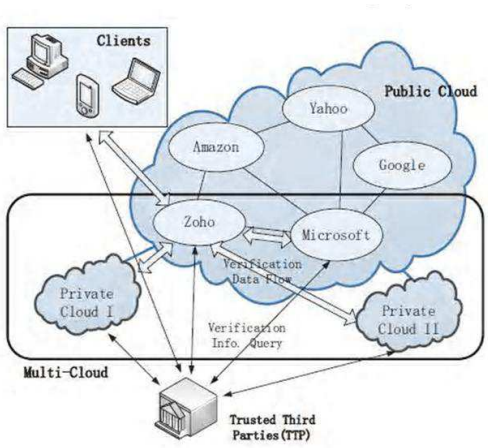


Fig 1: Verification architecture for data integrity

In this design, we have a tendency to contemplate the existence of multiple CSPs to cooperatively store and maintain the clients' information. Moreover, a cooperative PDP is used to verify the integrity and availableness of their keep information altogether CSPs. The verification procedure is delineated as follows: first of all, a shopper (data owner) uses the secret key to pre-process a file that consists of a group of _blocks, generates a set of public verification info that's keep in TTP, transmits the file and a few verification tags to CSPs,

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 8, August 2014)

and should delete its native copy; Then, by employing a verification protocol, the shoppers will issue a challenge for one CSP to envision the integrity and availability of outsourced information with relation to public info keep in TTP. We neither assume that CSP is trust to ensure the safety of the keep data, nor assume that information owner has the power to gather the proof of the CSP's fault when errors are found. to realize this goal, a TTP server is made as a core trust base on the cloud for the sake of security. we have a tendency to assume the TTP is reliable and freelance through the subsequent functions [12]: to setup and maintain the CPDP cryptosystem; to come up with and store information owner's public key; and to store the 9 public parameters accustomed execute the verification protocol within the CPDP theme. Note that the TTP isn't directly concerned within the CPDP theme so as to cut back the complexity of cryptosystem.

1.2 Existing System

There exist varied tools and technologies for multicloud, like Platform VM Orchestrator, VMware, v Sphere, and Overt. These tools facilitate cloud providers construct a distributed cloud storage platform for managing clients' information. However, if such an important platform is prone to security attacks, it would bring irrecoverable losses to the clients. for instance, the confidential information in Associate in Nursing enterprise could also be illicitly accessed through are mote interface provided by a multi-cloud, or relevant information and archives could also be lost or tampered with after they area unit keep into an · unsure storage pool outside the eEnterprise's therefore, it's indispensable for cloud service suppliers to produce security techniques for managing their storage services.

1.3 projected System

To check the supply and integrity of outsourced information in cloud storages, researchers have projected 2 basic approaches referred to as obvious information Possession and Proofs of Irretrievability .Ateniese et al. 1st projected the PDP model for guaranteeing possession of files on un trusted storages Associate in Nursing provided an RSA-based theme for a static case that achieves the communication value. They conjointly projected a publically verifiable version, that permits anyone, not simply the owner, to challenge the server for information possession..They projected a light-weight PDP theme supported scientific discipline hash perform and interchangeable key cryptography, but the servers will deceive the house owners by mistreatment previous data or responses as a result of the dearth of randomness within the challenges. The numbers of updates and challenges area unit restricted and glued in advance and users cannot perform block insertions anyplace.

1.4 Definition of Cooperative PDP

In order to prove the integrity of information keep in an exceedingly multi-cloud setting, we define a framework for CPDP supported interactive proof system (IPS) and multiprover.

1.5 HASH INDEX HIERARCHY FOR CPDP

To support distributed cloud storage, we have a tendency to illustrate a representative design used in our cooperative PDP theme as shown in Figure a pair of. Our design features a hierarchy structure that resembles a natural representation of file storage. This data structure $\mathcal{H}$ consists of 3 layers to represent relationships among all blocks for keep resources. they're delineated as follows:

1) categorical Layer: offers Associate in Nursing abstract illustration of the keep resources;
2) Service Layer: offers and manages cloud storage services; and
3) Storage Layer: realizes information storage on several physical devices.

We create use of this easy hierarchy to arrange information blocks from multiple CSP services into a largesize file by shading their variations among these cloud storage systems. for instance, in Figure a pair of the Resource in categorical Layer area unit split and stored into 3 CSPs, that area unit indicated by totally different colours, in commission Layer. In turn, each CSP fragments and stores the allotted information into the storage servers in Storage Layer. We have a tendency to conjointly create use of colours {to distinguish|to totally differentiate|to tell apart} different CSPs. Moreover, we follow the logical order of the info blocks to arrange the Storage Layer. This design also provides special functions for information storage and management, e.g., there may 12 exist overlaps among information blocks (as shown in dotted boxes) and discontinuous blocks however these functions could increase the quality of storage management

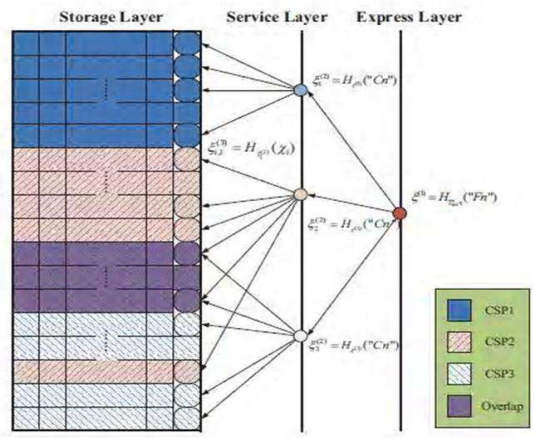


Fig.2: Index-hash hierarchy of CPDP model.

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 8, August 2014)

In storage layer, we have a tendency to outline a standard fragmentstructure that gives probabilistic verification of data integrity for outsourced storage. The fragmentstructure could be a organization that maintains a collection of block-tag pairs, allowing searches, checks and updatesin (1) time.

2. IMPLEMENTATION

Implementation is that the stage of the project once the theoretical style is clad into a operating system. So it is often thought of to be the foremost vital stage in achieving a successful new system and in giving the user, confidence that the new system can work and be effective. The implementation stage involves careful coming up with, investigation of the present system and it's constraints on implementation, planning of strategies to realize shift and analysis of shift strategies.

2.1 MODULES

- Multi cloud storages
- Cooperative PDP
- Data Integrity
- Third Party Auditor
- Cloud User

2.1.1 Multi cloud storages

Distributed computing is employed to talk to any giant collaboration within which many individual pc house owners permit a number of their computer's process time to be place at the service of an outsized downside. In our system the every cloud admin consist of information blocks .the cloud user transfer the info into multi cloud. Cloud computing atmosphere is made supported open architectures and interfaces, it has the aptitude to include multiple internal and/or external cloud services together to produce high ability. we have a tendency to decision such a distributed cloud environment as a multi-Cloud .A multi-cloud permits purchasers to simply access his/her resources remotely through interfaces.16

2.1.2 Cooperative PDP

Cooperative PDP (CPDP) schemes adopting zero-knowledge property and three-layered index hierarchy, severally. specially economical technique for choosing the best variety of sectors in every block to reduce the computation prices of clients and storage service suppliers. Cooperative PDP (CPDP) theme while not compromising information privacy supported trendy cryptographical techniques.

2.1.3 Information Integrity

Data Integrity is extremely necessary in info operations specially and information warehousing and Business intelligence normally. as a result of information Integrity ensured thatdata is of top quality, correct, consisten t and accessible.

2.1.4 Third Party Auditor

Trusted Third Party (TTP) whois trusty to store verification parameters and offerpublic question services for these parameters. In our system the trusty Third Party, view the user information blocks and uploaded to the distributed cloud. In distributed cloud environment every cloud has user information blocks. If any modification tried by cloud owner a alert is send to the trusty Third Party.

2.1.5 Cloud User

The Cloud Userwho encompasses a great deal of knowledge to be keep in multiple clouds and have the permissions to access and manipulate keep information. The User's information is converted into information blocks. the info block is uploaded to the cloud. The TPA views the data blocks and Uploaded in multi cloud. The user will update the uploaded information. If the user needs to transfer their files, the data's in multi cloud is integrated and downloaded.

3. LITERATURE SURVEY AND THEORETICAL TECHNIQUE

Literature survey is that the most vital step in code development method. Before developing the tool it's necessary to see the time issue, economy n company strength. Once these items square measure happy, 10 next steps square measure to see that software and language are often used for developing the tool. Once the programmers begin building the tool the programmers want ton of external support. This support are often obtained from senior programmers, from book or from websites. Before building the system the on top of consideration r taken thought into consideration} for developing the projected system.

4. DATA WORDBOOK

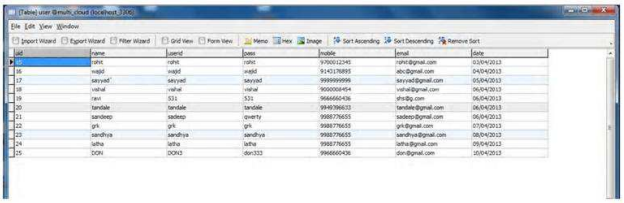
The logical characteristics of current systems information stores, together with name, description, aliases, contents, and organization. Identifies processes wherever the info are used and wherever immediate access to info required. is the idea for identifying info necessities throughout system style. Uses of knowledge wordbook

i. To manage the detail in giant systems
ii. to speak a standard that means for all system components
iii. To Document the options of the system
iv. To facilitate analysis of the main points so as to guage characteristics and confirm wherever system changes ought to be created.
v. To find errors and omissions within the systems

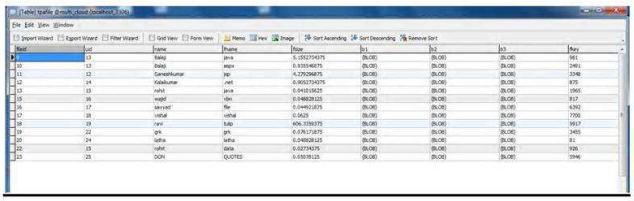
International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 8, August 2014)

5. DATABASE TABLES

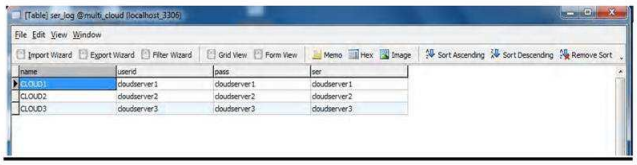
The data entered in the registration form are saved in this table.



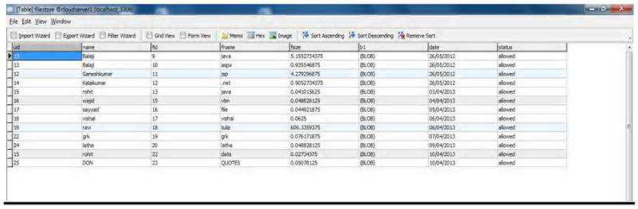
2. The files which are uploaded in the Cloud are saved in this table.



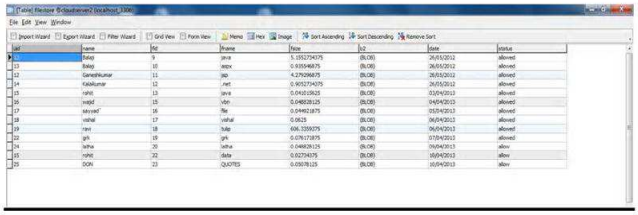
3. These are the username passwords for the three cloud servers.



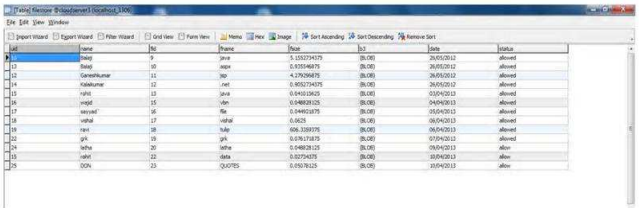
4. These are the files saved in the cloud server 1.



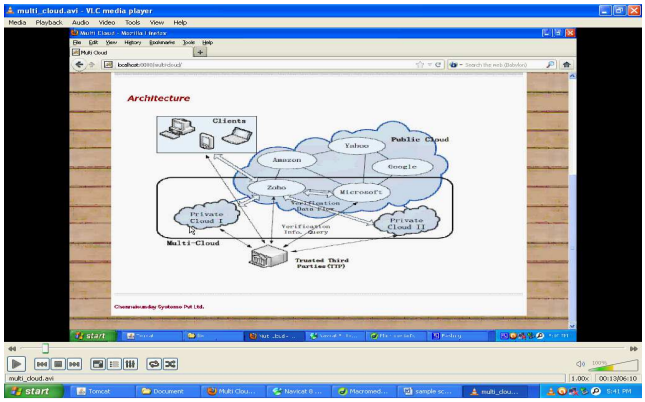
5.These are the files saved in the cloud server 2.

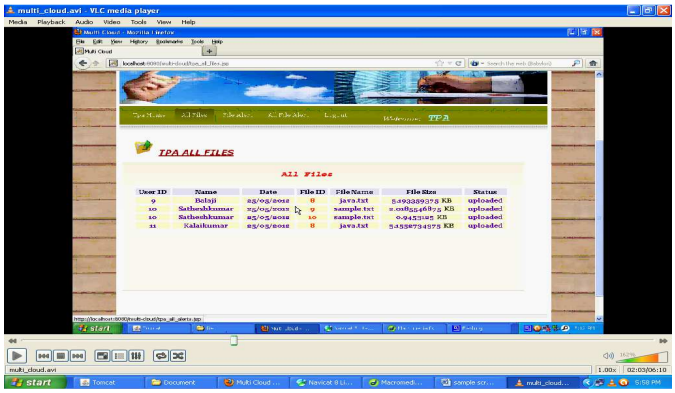


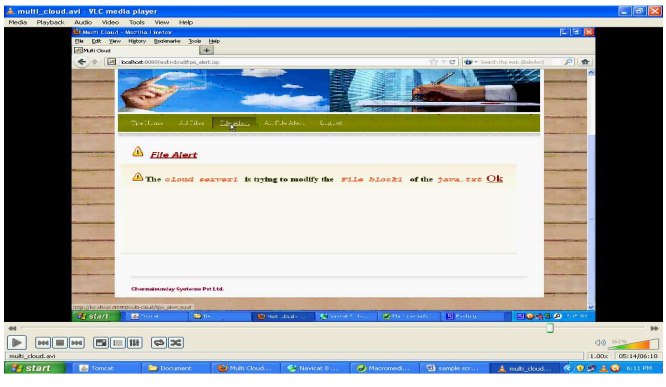
6.These are the files saved in the cloud server 2.



6. RESULTS







International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 8, August 2014)

6. CONCLUSION

we presented the construction of an efficient PDP scheme for distributed cloud storage. Based on homomorphic 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. 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]. B. Sotomayor, R. S. Montero, I. M. Llorente, and I. T. Foster, Virtual infrastructure management in private and hybrid clouds," IEEE Internet Computing, vol. 13, no. 5, pp. 14–22,2009.
[2]. G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner, Z. N. J. Peterson, and D. X. Song, "Provable data possession at untrusted stores," in ACM Conference on Computer and Communications Security, P. Ning, S. D. C. di Vimercati, and P. F. Syverson, Eds. ACM, 2007, pp. 598–609.
[3]. A. Juels and B. S. K. Jr., "Pors: proofs of retrievability for large files," in ACMConference on Computer and Communications
[4]. Security, P. Ning, S. D. C. di Vimercati, and P. F. Syverson, Eds. ACM, 2007, pp. 584–597.G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in Proceedings of the 4th international conference on Security and privacy in communication netowrks, SecureComm, 2008, pp. 1–10.
[5]. C. C. Erway, A. K¨upc¸ ¨u, C. Papamanthou, and R. Tamassia, "Dynamic provable data possession," in ACM Conference on Computer and Communications Security, E. Al-Shaer, S. Jha, and A. D. Keromytis, Eds. ACM, 2009, pp. 213–222.
[6]. H. Shacham and B. Waters, "Compact proofs of retrievability," in ASIACRYPT, ser. Lecture Notes in Computer Science, J. Pieprzyk, Ed., vol. 5350. Springer, 2008, pp. 90–107.
[7]. Q. Wang, C.Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in ESORICS, ser. Lecture Notes in Computer Science, M. Backes and P. Ning, Eds., vol. 5789. Springer, 2009, pp. 355–370.
[8]. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S. Yau, "Dynamicaudit services for integrity verification of outsourced storages in clouds," in SAC, W. C. Chu, W. E. Wong, M. J.Palakal, and C.-C. Hung, Eds. ACM, 2011, pp. 1550–1557.
[9]. K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availabilityand integrity layer for cloud storage," in ACM Conference on Computer and Communications Security, E. Al-Shaer, S. Jha, andA. D. Keromytis, Eds. ACM, 2009, pp. 187–198.
[10]. Y. Dodis, S. P. Vadhan, and D. Wichs, "Proofs of retrievabilityvia hardness amplification," in TCC, ser. Lecture Notes in Computer Science, O. Reingold, Ed., vol. 5444. Springer, 2009,pp. 109–127.
[11]. L. Fortnow, J. Rompel, and M. Sipser, "On the power of multiproverinteractive protocols," in Theoretical Computer Science, 1988, pp. 156–161.
[12]. Y. Zhu, H. Hu, G.-J. Ahn, Y. Han, and S. Chen, "Collaborativeintegrity verification in hybrid clouds," in IEEE Conference onthe 7th International Conference on Collaborative Computing: Networking,Applications and Worksharing, CollaborateCom, Orlando,Florida, USA, October 15-18, 2011, pp. 197–206.
[13]. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz,A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing,"EECS Department, University of California, Berkeley, Tech. Rep., Feb 2009.
[14]. D. Boneh and M. Franklin, "Identity-based encryption from the weil pairing," in Advances in Cryptology (CRYPTO'2001), vol. 2139 of LNCS, 2001, pp. 213–229.
[15]. O. Goldreich, Foundations of Cryptography: Basic Tools. Cambridge University Press, 2001.

About the authors:

Nagaraj Peddarapu currently working as a Associate Professor in CSE dept, Sri Indu College of Engg & Technology. He gained 6 years in teaching. His research interests include: Data Mining, Cloud Computing and Data Base Management Systems.

Dr. Ch G.V.N. Prasad currently working as a Professor and HOD of CSE dept, Sri Indu College of Engg & Technology. He gained 12 years of experience in IT industry (8 years in National Informatics Centre, Govt. of India, as Scientist and Software Analyst in AT&T in US) and 11 years of experience in Teaching (As a Professor & HOD of CSE Dept).