RESEARCH ARTICLE OPEN ACCESS

Hash Function based Data Partitioning in Cloud Computing for Secured Cloud Storage

*Parisha[1],Pooja Khanna[2],Puneet Sharma[3],Sheenu Rizvi[4]
Department of Computer Science & Engineering, Amity University, Lucknow
*Corresponding author: *Parisha*

ABSTRACT
There is a necessity to firmly manage, store, and examine enormous quantity of complex data i.e. semi-structure and unstructured to regulate configurations and developments in order to expand the feature and excellence. It is very essential that clouds be secure and protected because of the critical environment of the applications. The foremost security task using cloud is that the holder of the data possibly will not have control of where the data is located. If anyone who desires to exploit the advantages of using cloud computing, one must also use the resource sharing and scheduling delivered by clouds. The data stored on cloud may be possible altered or modified without the knowledge of client. There should be a mechanism that verifies the stored and the data being retrieved is same.Thus we need protection of the data.This paper examines some significant security services including verification, encryption and decryption are provided in cloud computing system.

In this paper data partitioning technique is used for privacy conserving and security of data, using third party auditor (TPA) and concept of hash function. A hash function takes a data of variable size and generates a data of fixed length, which is unique for each data. when data is encoded the hash value of encrypted data is also created and it is stored by third party auditor .For retrieval of data, third party auditor decodes the data by random key and creates the hash of the encrypted file attained from cloud then the hash value is matched and compared with one generated by it and the hash value at data holder, this verifies the data exactness. The objective of this work is designing, an effective flexible storage structure to confirm the accessibility of data and data exactness in cloud.

***Keywords*:** Authentication, Cloud computing,Encryption and decryption,Hash function, Integrity

Date of Submission: 10 -07-2017 Date of acceptance: 31-07-2017

I. INTRODUCTION

Cloud computing considerably improves team work, agility, and scale, therefore, it allowing a worldwide computing scheme over the Internet. The key features of cloud computing that include resource combining, fast elasticity, wide web access, measured facilities, just single click away, easy usage, just pay for the services you use and location independent. All of these makes using cloud flawlessly and transparently [1] [2]. Cloud computing is a technology that is based on internet in which applications, files and various resources are easily retrieved and pooled by the users over the web in efficient and flexible manner. Cloud computing runs service models which are-Platform as a service (PaaS), Software as a service

(SaaS) and Infrastructure as a service (IaaS) and the deployment models are-private cloud, public cloud and hybrid cloud. The benefits of cloud computing makes it more popular which are on request service, wide web access, measured services, just single click away ,easy usage, just pay for the services you use and location independent. These advantages also rise the security threat to data which has to be coped. The network based connections which are reliable and their bandwidth makes it easier and possible that users may access high feature facilities from data. It reduces the duty of local systems for data preservation identically. As an outcome, workers

are dependent on their cloud service providers (CSP) for data integrity and accessibility of data [3] [4]. Even though the cloud setups are much influential and trustworthy than particular devices, still exist the extensive variety of both external and internal risk for integrity of data. Instances of data damage occurrences of notable storage of cloud services seem time to time. Meanwhile users cannot preserve a local replica of data, existence of numerous reasons for service provider to behave faithlessly toward the users concerning the position of the outsourced information [5] [6]. For example, by reducing cost to increase the profit edge, CSP can discard infrequently accessed data without being noticed in appropriate manner. Likewise, CSP can effort to hide data damage occurrences to uphold a status. So, while outsourcing the data in the cloud is financially striking for the price and difficulty in data storage, it's deficient of allowing strong assertion of integrity and accessibility can hinder its extensive acceptance by distinct cloud users [7].

In the direction of attaining the assurance of cloud information integrity, accessibility and impose are the features of cloud storing service.

On behalf of cloud user, well-organized procedures that allow on request data accuracy authentication to be designed [8] [9] [10].Though, the point that users may not have extensive physical resistor of information that excludes the straight acceptance of primitives for the need of data integrity security. The data that is kept in the cloud can be effortlessly retrieved and also modified by the users. Therefore, it is authoritative to support this additional quality in the storage of cloud accuracy, by which the storage scheme design become further challenging [11]. Cloud computing offers right to use the data but the challenges of cloud is to confirm that approved individuals may only access to it. While routine cloud settings we depend on the third parties to give verdicts related to data and schemes in various methods which has been never seen previously in cloud computing. It is very difficult to have suitable schemes to avoid the providers of cloud consuming client's data in a manner which has not got permitted.

Placing data in the cloud allows excessive suitability to the consumers as need not to worry regarding the difficulties in managing hardware and execution of the technical information. Cloud storage is a service for developers and client. Developers are to access and store data in cloud and client can access the cloud by using client devices. The resources will be managed by cloud service provider. Flexibility, increased adeptness , scalability, capital overheads and reduced cost are the benefits of cloud storage and also they able to remove the data damage hazard [12].In recent times many work focus in the direction of third party auditing and remote data integrity testing. Data integrity testing at unreliable servers is the major concern with cloud data storage. Another concern is supportive data procedure for storage of cloud requests. To solve the problem of data integrity testing, many outlines are suggested which are-unbounded use of queries, high scheme efficiency and stateless verification etc. [13]. Data robustness is also an important necessity for effective storage systems. Various suggestions are there for storing data on servers. A method to maintain data robustness is to reproduce a data so that the replica of the data is stored by every single server and message can be retrieved easily as long as the storage server lasts [14].

II. LITERATURE SURVEY

In this section we did the literature survey related to data partitioning technique for security of data using third party auditor (TPA) for Safe and protected cloud data storage and privacy conserving as well as data integrity checking, storage of data and dynamic data storage with token pre-computation and how it is stored in the cloud is look over and analyzed [15].This effort studies the problem of confirming the dataintegrity storage in cloud and the job of permitting third party auditor (TPA) on the cloud to validate the integrity of the dynamic data stored in the cloud system. TPA can eliminate the participation of the client through inspecting whether his data that is stored in the cloud are absolutely complete or not [16] [17]. The idea of Third Party Auditor is manager and checker who checks between the cloud data storage server and machine. Third party auditor works on the basis of two categories that is public and private auditability. Public inspection permits anyone, not only approve client but also give the ultimatum to cloud server though storing none of secretive information for the purpose of exactness of storage data, whereas private auditability delivers greater efficiency to the client who are authorized. To decrease the effort of data management of the client the third party auditor audits the data of client. The auditor has the rights for attaining financial prudence for cloud computing, it terminates the link of the user by checking the information which is warehoused in the cloud are essentially together. The audit report according to the audit which can help client to compute the threat of the cloud information resources and also it is useful for service provider to expand and recover cloud service policy. Later TPA give the conformation to client to assure that his data is safe and secure in the cloud and the overall managing data is easy and not as much of challenging to client .Clouds has no limitations so that information may be placed wherever. This feature of cloud rise different issues interrelated to user authentication and data Confidentiality [18], [19].

Partitioning of data in horizontal and vertical order as discussed in [17, 18]. Here the data is partitioned into buckets and slicing technique is used for data storage onto cloud. In the works, [15], inventor deliberates producing signature methods for confirming the cloud data storage security. This method offers security of data storage. KiranGabhale [12] [20] define an outline on the Partitioning of data Technique that is done for data integrity testing and storage of data scheme that are presently used in dynamic multi transactional submission. By the dynamic storage of data with pre computation token and algorithms like AES, it examines that how it is efficiently stored in cloud. Integrity testing idea is used to discover and detect the server which is not behaving in proper manner considering data rectification and error exposure. To attain the quality of data, integrity, accessibility of trustworthy storing services, distributed structure is used and to perform these operation the data storage scheme with dynamic data operation method is applied. RSA is used to encrypt the data for security analysis. It contains a private and public key. The public key is known to one and all and it is considered for encryption of messages.

The elementary idea of this process are generating key for encoding and decoding of message. For security purpose encryption method is used to encode the folders and files by creating cipher and key producer object. Then, by initializing with private key, secret key is produced using cipher object. Decryption method is used to decode the folders and files and private key is created to retrieve and access files from the cloud. Separate key is generated for every end user to access the files from any position with security. Non- shared key is used for file decryption. After that the partitioning of data is done in alphabetic manner with the use of index technique.in this process firstly it checks in the folder that, first two letter which is retrieved having same letter or not. If the letter does not exist in the folder then create a folder and store that file in that particular folder. Then encrypt all partition files with public key and decrypt original file with private key when need to access. Ensuring secured data in cloud storage distributed storage scheme is also used. The analysis of data integrity in cloud storage is done in research work. Public audit ability and dynamic data operation are used to associate the integrity of data. The main idea of this effort is to obtain self-determining perception and service quality estimating with the third party auditor. To improve efficiency Storage scheme is planned to support multiple inspecting jobs.

In the work [21] focuses on data integrity testing and storage of data scheme that are presently used in dynamic multi transactional submission. The dynamic data storage offer facts regarding current storage scheme. Data partitioning is done in horizontal and vertical directions as Discussed here. The data is partitioned into number of buckets and after that slicing method is used for storage of data [23], [22]. In this work author reflects producing signature approaches for confirming the security of cloud storage. By using the RSA method dynamic operations are supported. Integrity of data and their exactness that is stored in cloud is discussed in this method. To identify the accessibility of information error rectification and data integrity inspection is used in cloud. Data error recovery and availability of data mechanisms are not given much importance. Symmetric key cryptography for security of storage and availability is discussed in [23] with partitioning scheme.

In the survey, ample of the debates and discussions are linked to the works, that certifies to contain copy of data in the local system. Particular drawback is reduced with the planned method system in [24]. Token pre-computation token process confirms data operation dynamically. It offers security to data storage system [25]. The drawback of this scheme is that to execute the processing of data encoding and decoding methods for security of data stored in cloud, it consumes extra cost and time. Partitioning technique plays significant part in this work. When there is a need arises to access the data the larger files breaks up into smaller parts for storing the data efficiently. There is much trouble in storing the data in cloud due to the complexity of data, thus to make it easy partitioning function is used in cloud. The files which are partitioned are encrypted along with Public key and then those data will be stored in cloud, while the data is served for storing in cloud partitioning takes place automatically. When there is a necessity to retrieve the same data, original data is also rebuilt [26].

III. PROPOSED SYSTEM

Here a structure is suggested to offer secure storage of data in cloud computing. This proposed system uses the partitioning technique along with the concept of hash function.

3.1 RSA:

RSA (Rivest-Shamir-Adleman) is an algorithm used by recent computers for encryption and decryption of data .This algorithm includes a public key and private key. The public key can be known to everyone and it is used to encode data .Data which are encoded with public key only decoded with private key .RSA can be used for generating protected communications network, confirmation and the identification of cloud service provider. In proposed system RSA is used to discover out the key pair for both user and third party auditor.These keys are used to encode and decode the file.

3.2 Hash Function:

A hash function takes a data of variable length and produces a data of fixed length. It produces small and static length data which is unique for each data. The hash code is also specified as message digest or Hash Value. It is a task of all the bits of the message and it delivers fault exposure ability. Any kind of change to any bits in the data consequences in a huge alteration to the hash code. The foremost necessities for the safety of hash functions are that they essential be one way functions and be crash unaffected. At this point, in the suggested structure the hash value of the file is considered and calculated as a result that integrity of data can be preserved.

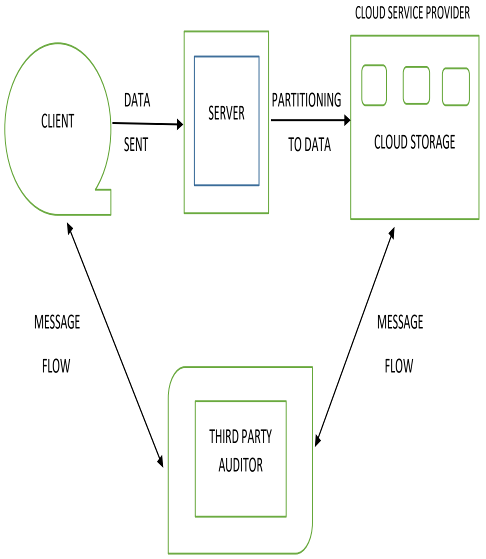
3.3 Participants involved:

In proposed system the participants which are involved are- user, third party auditor and cloud service provider.

1) User: It is an individual who has huge data to upload, store, access and retrieve from cloud storage and dependent on cloud system for storage of data and their computation and consumes the storage services offered by cloud service provider.

2) Third Party Auditor: It is a TPA, who is expert in his work and has capabilities and many responsibilities that users may not have. Third party auditor is the mediator between the Owner and cloud service provider who checks the data integrity stored on cloud.

3) Cloud Service Provider: cloud service provider manage all the important data storage services and provides storage space and has computation resources and offers storage services to the users. The user act together with cloud servers through service provider to access his data.



Figur**e 1:** Data integrity checking using third party auditor.

3.4 The Proposed Scheme involves following Operations:

1) Key Generation: user uses RSA algorithm for generation of blend of public and private key. The third party auditor also uses RSA for producing key pair for its own.

2) Key Sharing: The key sharing among data holder and third party auditor. Here only public key of third party auditor is swapped among data holder using source channel.

3) Data uploading: When user wants to store his data, user uses RSA for generation of blend of public and private key. Then the encryption of data to be stored is done by data owner. Now, the hash of encrypted data is generated. At the same time the overall hash value and encrypted data is re-encrypted with public key of third party auditor. Third party auditor generates a random key for executing encryption on the data generated after encryption. That random key generated by third party auditor is stored for performing decryption in future. After that, encrypted data is sent to server by the user. At server side the data is partitioned, the larger files break up into smaller parts for storing the data efficiently and the outcome is sent to cloud for storage.

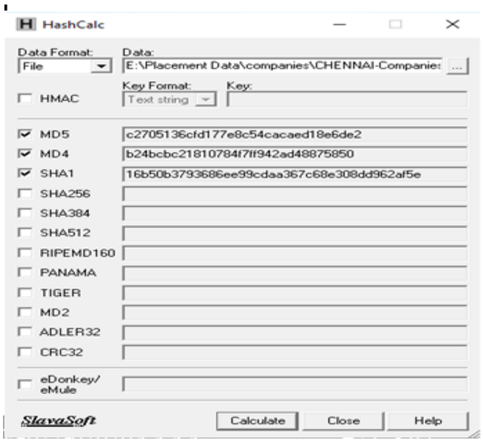
4) Data Retrieval:
When the owner of data request for the files stored in Cloud. The third party Auditor initially decodes the data by random key. Then the third party auditor produces the hash of the encrypted file that is acquired from cloud. Further, third party auditor compares the hash value with one produced by it and the hash value at data holder. The third party auditor confirms the correctness of data. Conferring to result obtained, third party auditor sends requested file to user specifying the correctness or not. After receiving encoded folder or file, user decode it by private key of him.

3.5 Case study

Several issues are considered related to security in cloud and many methods are applied to secure the data on cloud by cloud service providers. Although, these methods do not confirms the data correctness and integrity of data. Here, to check the data, proposed system uses the partitioning technique along with the concept of hash function to maintain the integrity of data and data correctness. Third party auditor (TPA) can check the data correctness on user's request.

Here, we are using an application termed HashCalc to calculate the hash value of file including documents, video games, movies, software, music, and others. Third party auditor can choose the file to be calculated and the desired hash functions from nine hash functions including MD4, MD5, SHA1, SHA256, MD2, CRC32 and more.

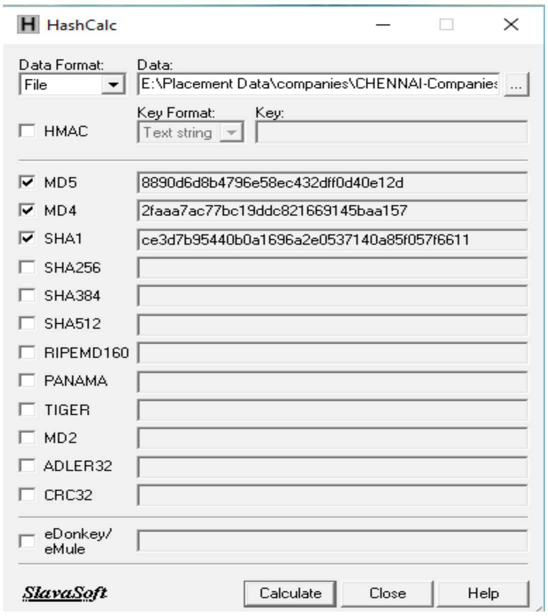
It can calculate for large-sized files up to 15 GB. After selection of hash function, TPA can specify the key format (hex string or text string) and the key. Now, we take a file and upload it in the application and calculate MD4, MD5 and SHA1 from the given hash functions. The resultant hash values of that file are represented as Fig 2:



Data Format: File
Data: E:\Placement Data\companies\CHENNAI-Companie:
□ HMAC Key Format: Text string Key:
☑ MD5 c2705136cfd177e8c54cacaed18e6de2
☑ MD4 b24bcbc21810784f7ff942ad48875850
☑ SHA1 16b50b3793686ee99cdaa367c68e308dd962af5e
□ SHA256
□ SHA384
□ SHA512
□ RIPEMD160
□ PANAMA
□ TIGER
□ MD2
□ ADLER32
□ CRC32
□ eDonkey/ eMule

Figure 2. Pre-computed hash value

After calculating the hash value of the file .we modify some data of that same file and again the hash value of modified file is calculated by uploading it in application. The calculated hash values are:



Data Format: File
Data: E:\Placement Data\companies\CHENNAI-Companies

☐ HMAC Key Format: Text string Key:

☑ MD5 8890d6d8b4796e58ec432dff0d40e12d
☑ MD4 2faaa7ac77bc19ddc821669145baa157
☑ SHA1 ce3d7b95440b0a1696a2e0537140a85f057f6611
☐ SHA256
☐ SHA384
☐ SHA512
☐ RIPEMD160
☐ PANAMA
☐ TIGER
☐ MD2
☐ ADLER32
☐ CRC32
☐ eDonkey/ eMule

Figure 3. Re-computed hash value

Above calculation shows that the resultant hash value is different after modification of data within same file.This confirms that data has been compromised. So, this concept can be applied on cloud environment to maintain the integrity of data and data accuracy.

IV. CONCLUSION AND FUTURE WORK

The proposed scheme delivers a safety and security scheme for securing the data in cloud computing with the help of hash function to deliver privacy and integrity to the data stored in cloud. This work consumes the method of partitioning the data in well-organized way; it also deals with high storage size along with less reduction of time. Encryption and decryption technique is used in our system. These are basic steps which are used to maintain the security of the data. The key benefit of this system holds the safety and security of data along with preserving the integrity of the data. The process of partitioning permits storing of the data in simple and efficient way. It also offers less cost in storage of data, a method for flexible data retrieval and the time and space is proficiently reduced during storage. Dynamic operation is another idea in which encoding and decoding method makes the data safe and secure even though storing in cloud and the data integrity checking avoids all the threats and servers which are not behaving well ensuring data security. Data integrity testing also classifies the threats and misbehaving server while storing the data in cloud confirming data security. The foremost benefit of this technique that it includes the security and protection of data along with keeping the integrity of the data conserved. The schemes which are different and related to scattering of encoding and decoding key to retrieve the data from server. We can conclude that in what manner the data can be proficiently and securely accessed from the server completely subjected to the proportion of the secret key that is the reduced size of the secret key the minor necessity of space. Further, forthcoming work is scheduled to offer advanced way of security, storage space and examining schemes for computation of outsourced data in cloud.

REFERENCES

[1]. Sun Microsystems, Inc., “Building Customer Trust in Cloud Computing with Transparent Security,” https://www.sun.com/offers/details/sun_transparency.xml, Nov. 2009.

[2]. K. Ren, C. Wang, and Q. Wang, “Security Challenges for the Public Cloud,” IEEE Internet Computing, vol. 16, no. 1, pp. 69-73, 2012.

[3]. Sun Microsystems, Inc., “Building Customer Trust in Cloud Computing with Transparent Security,” https://www.sun.com/offers/details/sun_transparency.xml, Nov. 2009.

[4]. K. Ren, C. Wang, and Q. Wang, “Security Challenges for the Public Cloud,” IEEE Internet Computing, vol. 16, no. 1, pp. 69-73,2012.

[5]. E. Bertino, F. Paci, and R. Ferrini, “Privacy-Preserving Digital Identity Management for Cloud Computing,” *IEEE* Computer Society Data Engineering Bulletin, Mar. 2009, pp. 1–4.

[6]. J. Kincaid, “MediaMax/TheLinkup Closes Its Doors,” http://www.techcrunch.com/2008/07/10/mediamaxthelinkup-closesits-doors, July 2008.

[7]. B. Krebs, “Payment Processor Breach May Be Largest Ever,”http://voices.washingtonpost.com/securityfix/2009/01/payment_processor_breach_may_b.html, Jan. 2009.

[8]. A. Juels and B.S. Kaliski Jr., “PORs: Proofs of Retrievability for Large Files,” Proc. 14th ACM Conf. Computer and Comm. Security (CCS ’07), pp. 584-597, Oct. 2007.

[9]. G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, “Provable Data Possession at Untrusted Stores,” Proc. 14th ACM Conf. Computer and Comm. Security (CCS’07), pp. 598-609, Oct. 2007.

[10]. M.A. Shah, M. Baker, J.C. Mogul, and R. Swaminathan, “Auditing to Keep Online Storage Services Honest,” Proc. 11th USENIX Workshop Hot Topics in Operating Systems (HotOS ’07), pp. 1-6, 2007.

[11]. J. Kubiatowicz, D. Bindel, Y. Chen, P. Eaton, D. Geels, R.Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao, “Oceanstore: An Architecture for Global-Scale Persistent Storage,” Proc. Ninth Int’l Conf. Architectural Support for Programming Languages and Operating Systems (ASPLOS), pp. 190-201, 2000.

[12]. SwapnilV.Khedkar,A.D.Gawande, “Data Partitioning Technique to Improve CloudData Storage Security.” IJCSIT, Vol. 5 (3), 2014.

[13]. M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu,“Plutus: Scalable Secure File Sharing on Untrusted Storage,” Proc.Second USENIX Conf. File and Storage Technologies (FAST), pp. 29-42, 2003.

[14]. Hsiao-Ying Lin; Tzeng, W.-G.; , "A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding," Parallel and Distributed Systems, IEEE Transactions on , vol.23, no.6, pp.995-1003, June 2012.

[15]. Slicing: A New Approach for Privacy Preserving Data Publishing Tiancheng Li ; Ninghui Li ; Jian Zhang ; Molloy, I. Knowledge and Data Engineering, IEEE Transactions on Volume: 24, Issue: 3 DOI: 10.1109/TKDE.2010.236 Publication Year: 2012, Page(s): 561 – 574.

[16]. Toward Secure and Dependable Storage Services in Cloud Computing Cong Wang ; Qian Wang ; KuiRen ; Ning Cao ; WenjingLouServices Computing, IEEE Transactions on Volume:5DOI:10.1109/TSC.2011.24 PublicationYear:2012,Page(s):220-232.

[17]. PDDS - Improving cloud data storage security using data partitioning technique Selvakumar, C. ; Rathanam, G.J. ; Sumalatha, M.R. Advance Computing Conference (IACC), 2013 IEEE 3rdInternational DOI: 10.1109/IAdCC.2013.6506806 Publication Year: 2013.

[18]. Hsiao-Ying Lin; Tzeng, W.-G.; "A Secure Erasure Code- Based Cloud Storage System with Secure Data Forwarding," Parallel and Distributed Systems, IEEE Transactions on, vol.23, no.6, pp.995-1003, June 2012.

[19]. C. Wang, Q. Wang, K. Ren, and W. Lou, “Ensuring Data Storage Security in Cloud Computing,” Proc. 17th Int’l Workshop Quality of Service (IWQoS ’09), pp. 1-9, July 2009.

[20]. Kiran Gabhal1, NarendraJadyal, Anurag More, VinayakBhalekar, Prof. V.V. Dakhode5, “ Data Partitioning Technique to Improve Cloud Data Storage Security”,International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, March 2016.

[21]. C. Selvakumar, G. JeevaRathanam, M. R. Sumalatha, “PDDS - Improving Cloud Data Storage SecurityUsing Data Partitioning Technique”, 2013 3rd IEEE International Advance Computing Conference (IACC).

[22]. Tiancheng Li; Ninghui Li; Jian Zhang; Molloy, I.; "Slicing: A New Approach for Privacy Preserving Data Publishing," Knowledge and Data Engineering, IEEE Transactions on, vol.24, no.3, pp.561-574, March 2012.

[23]. Paredes, L.N.G.; Zorzo, S.D.;, "Privacy Mechanism for Applications in Cloud Computing," Latin America Transactions, IEEE (Revista IEEE America Latina) , vol.10, no.1, pp.1402-1407, Jan. 2012.

[24]. C. Wang, Q. Wang, K. Ren, and W. Lou, “Privacy-Preserving Public Auditing for Storage Security in Cloud Computing,” Proc. IEEE INFOCOM, Mar. 2010.

[25]. Shephali Singh, Puneet Sharma, Dr. Deepak Arora, “Secure Outsourcing of Linear Programming in Cloud Computing Environment: A Review,” IJERA, http://www.ijera.com/papers/Vol7_issue4/Part-6.

[26]. C. Wang, K. Ren, W. Lou, and J. Li, “Towards Publicly Auditable Secure Cloud Data Storage Services,” IEEE Network Magazine, vol. 24, no. 4, pp. 19-24, July/Aug.2010.

International Journal of Engineering Research and Applications (IJERA) is **UGC approved** Journal with Sl. No. 4525, Journal no. 47088. Indexed in Cross Ref, Index Copernicus (ICV 80.82), NASA, Ads, Researcher Id Thomson Reuters, DOAJ.

Parisha. "Hash Function based Data Partitioning in Cloud Computing for Secured Cloud Storage." International Journal of Engineering Research and Applications (IJERA) 7.7 (2017): 01-06.