RESEARCH ARTICLE OPEN ACCESS

Data Partitioning Technique In Cloud: A Survey On Limitation And Benefits

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

ABSTRACT
In recent years,increment in the growth and popularity of cloud services has lead the enterprises to an increase in the capability to handle, store and retrieve critical data. This technology access a shared group of configurable computing resources, which are- servers,storage and applications. Cloud computing is a succeeding generation architecture of IT enterprise, which convert the application software and database to large data hubs.Data security and storage of data is an essential functionality of cloud services.It allows data storage in the cloud server efficiently without any worry. Cloud services includes request service, wide web access, measured services, just single click away ,easy usage, just pay for the services you use and location independent.All these features poses many security challenges.The data partitioning techniques are used in literature, for privacy conserving and security of data, using third party auditor (TPA). Objective of the current workis to review all available partitioning technique in literature and analyze them. Through this work authors will compare and identify the limitations and benefits of the available and widely used partitioning techniques.

Keywords: Cloud Computing, Cloud Security, Cloud Storage,Partitioning Technique, Third Party Auditor.

I. INTRODUCTION

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[2]. 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[1][2].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 anoutcome, workersare 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 reasonsfor service provider to behavefaithlessly toward the users concerning the position of the outsourced information[5][6][7][8][9]. 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 [9].
To attain the assurance of cloud information integrity, accessibility and imposeare the features of cloud storing service. On behalf of cloud user, well-organized procedures that allow on request data accuracy authentication to be designed [10][11][12][13].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 [14][15]. 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.

Putting 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 [16].In recent times many work focus in the direction of third party auditing and remote data integrity testing. Data integrity testing at untrusted 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. [17]. 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 [18].

In cloud system the client put his data on cloud servers and it is locally maintained. Data partitioning technique offers security to data and by which user can avoid local copy of data or multiple copies of data. In cloud storage the architecture of network has different network entities as represented by Fig 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.

• **Cloud Server (CS):** Cloud server is an entity in which cloud service provider (CSP) manage all the important data storage services and provides storage space and has computation resources.

• **Third Party Auditor (TPA):** It is a TPA, who is expert in his work and has capabilities and many responsibilities that users may not have. TPA is a mediator between client and cloud data storage.TPA is trustworthy to evaluate and uncover hazard of cloud data storage services on user's request. In cloud data storage, user stores the data througha CSP into a number of sets of cloud servers in cloud data storage. The user act together with cloud servers through service provider to access his data. TPA (Third Party Auditor) executes partitioning of the data which is uploaded by the user on cloud. [19]

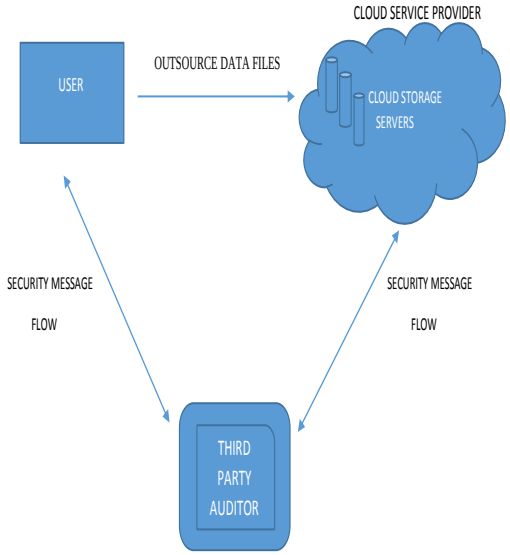

Figure 1: Data storage in cloud

II. LITERATURE SURVEY

In this section various data partitioning techniques that has been used by researches is analyzed. The data partitioning techniques are used for privacy conserving and security of data, using third party auditor (TPA). The focus of the study is to analyze commonly used comparative schemes and procedures available in literature for attaining data safety and confidentiality. [20]

Qian Wang[21] define an outline that allows third party auditor (TPA), on behalf of the cloud user for building of block tag validation, Merkle Hash Tree (MHT), is used. MHT is built as a binary treein which effective and secure verification of data is done. On the other hand it contains several disadvantages like third party can be dishonest or may not be capable to commit essential functioning outresources executing persistent validation. This model has three procedures, a digital signature portion will be done by the user then the cloud server validates the information that is stored in the cloud to verify the management and disturbance in the cloud data. Later the auditor confirms the cloud server to validate and verify the cloud server if the server was deploying in the data.

Cong Wang, Qian Wang, KuiRen, [1] define a way of distributed storage integrity auditing mechanism. It acquires certain benefits which are - dynamic data authentication and malicious data reform occurrence. But drawbacks

of this technique that are numerous replica of data exists which needs a large memory space so, Data Integrity is hard to attain.

Himika Parmar, Nancy Nainan [22] define an authorized service technique which are image based authentication (IBA) and one time password, which provide extraordinary way of security and confidentiality scheme.

Huaqun Wang [23] [24] describe proxy provable data possession scheme. It has certain benefits which are effective user organized management of data efficient storage. Drawbacks like Public validation can be a reason of intruder's crash on information files.It is examined that how the data storage along with pre computation token and advanced encryption standard (AES) algorithm is efficiently stored in the cloud server. Integrity testing concept is also used to identify and detect the server which is not behaving well considering data adjustment and error exposure.

C. Wang, Q. Wang [25] describes file distribution and token pre- computation system. But the problem is that it has no solution to offer the files block insertion operation. When block is inserted it requires lots of shifting of block.

Kiran Gabhale [26] focuses 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(Rivest-shamir-Adleman) 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 firstlyit 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 obtainself-determining perception and service quality estimating with the third party auditor. To improve efficiency storage scheme is planned to support multiple inspecting jobs.

C.selvakumar [27] 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 [28][29]. 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. Whereas symmetric key cryptography for security of storage and availability is discussed in [1] with partitioning scheme. Integrity of remote data testing has various stimulating issues in cloud storage services.

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 [30]. Token pre-computation process confirms data operation dynamically. Itoffers security to data storage system.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 [27].

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 is 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 [31] [32].

TPA uses the homomorphism authenticator and random mask technique to assure the user is not aware about the exact content of data that is warehoused safely on the server throughout the procedure of the checking, it eradicates the load of client as of the complex and costly inspecting job .It also make sure that the user is relieved from the fear of data leak. TPA can handle multiple audit sessions at the same time. It will be further extended into a multi-user set where TPA may accomplish the severa linspecting jobs in a fully batch method. A cloud-based capacity plot that permits the information holder to gain from the offices accessible by the service provider and build shared belief among them. The mechanism has fourcrucial peculiarities [33]:

(i) It permits the holder to outsource the information to a cloud service provider, and implement procedures on the outsourced information, i.e., piece alteration, Insert, removal, and attach.
(ii) It gives assurance that only permitted clients get best recent interpretation of the outsourced data.
(iii) It authorizes unplanned common belief between the manager and CSP.
(iv) It allows the manager to accept or reject all access to the outsource ddata.

The security issues in this plan are: The information has identical copies in the locality framework. This instrument gives information stockpiling security. The existing system takes extratime and expenditure to achieve the element altering of Information encryption and separating procedures to store information in cloud [19][34].In table 1,comparison of different schemes used by data partitioning technique in cloud data storage security along with the tools and method, its usage and the issues with that schemes.

Table 1: Comparison of different schemes used by data partitioning technique in cloud data storage security.

SCHEME	TOOLS AND METHOD	USE	ISSUE
The scheme of third party auditor (TPA). [21]	Merkle Hash Tree (MHT).	It is used for the construction of block tag validation.	The third party can be dishonest and may not be able to commit constant verification.
Data integrity and storage. [1][26]	Distributed storage integrity auditing mechanism.	Data verification and data modification.	Method possess multiple copies which requires large storage space.
Efficient data management. [23][24]	Proxy provable data possession method.	User controlled data management.	Intruder collision on data files.
Data integrity checking and storage mechanism. [26][27]	RSA algorithm and data partitioning technique.	To avoid misbehaving server and achieve integrity dependable storage services.	It consume extra time and cost to execute.
Authenticati on and identity.[22]	Image based authenticati on and one time password.	Used for high level security mechanism.	Increased complexities and faced problem when multiple cloud service providers are there.

III. CONCLUSION AND DISCUSSION

Partitioning technique provides a great advantage to a wide variety of requests by improving manageability, availability and performance. Partitioning is an important tool for building multi-terabyte structures or systems with tremendously high availability requirements. This technique can significantly reduce the total cost of data ownership by keeping older appropriate information still online on the low cost storage devices. Efficiency is another advantage of this technique in which the records which are used together are grouped together and each partition is optimized for performance, recovery and security, it also take the advantage of parallel processing capability and it offers detailed advantages to enhance resource utilization and reduces the execution time. Partitioning enables an effective, simple, and very powerful approach when considering information management for large environments. Having the advantage of load balancing reduces contention by storing partitions on different disks [27]. It enables information administration processes such as information loading, index formation and, backup and restoration. As an outcome the processes become more faster. This technique not only improves the management of the large data centers, it also permits the medium and smaller range databases to take preference of its benefits. The performance level gains enhancement because the queries can be solved easily [26].

This technique increases the administration overhead and it is more challenging to create a consistent snapshot of an application which is running on different bulks. The method of partitioning has more chances to waste disk space and there is a chance that makes incidents involving disk full. Partitioning of data consumes more resources and non-transparent partitioning increases complexities [27].

In this paper, efficient and effective data storage and security of data in cloud service is reviewed. The data partitioning technique allows storing of datain much easier and efficient way. It reduces the extra time, space and cost and gives an efficient way for flexible access of data. Dynamic operation is a key idea in which encryption and decryption procedure makes the data safe and secure while storing in cloud and the data integrity testing avoids all the threats and servers which are not behaving well ensuring data security.It utilizes the method of separating the data in a proficient way and also offers high storage capacity .The major benefit of this technique that it includes the security and protection of data along with keeping the integrity of the data conserved. The structures which are different and related to scattering of encryption and decryption key to access the data from server. We can determine 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]. 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.

[2]. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage", IEEE Transactions on computers, vol. 62, no. 2, february 2013.

[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]. M. Arrington, "Gmail Disaster: Reports of Mass Email Deletions,"http://www.techcrunch.com/2006/12/28/gmail-disasterreportsof-mass-email-deletions, Dec. 2006.

[6]. J. Kincaid, "MediaMax/TheLinkup Closes Its Doors," http://

[7]. www.techcrunch.com/2008/07/10/mediamaxthelinkup-closesits-doors, July 2008.

[8]. Amazon.com, "Amazon S3 Availability Event: July 20, 2008,"http://status.aws.amazon.com/s3-20080720.html, July 2008.

[9]. S. Wilson, "Appengine Outage," http://www.cio-weblog.com/50226711/appengine_outage.php, June 2008.

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

[11]. 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.

[12]. 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.

[13]. M.A. Shah, M. Baker, J.C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIXWorkshop Hot Topics in Operating Systems (HotOS '07), pp. 1-6, 2007.

[14]. M.A. Shah, R. Swaminathan, and M. Baker, "Privacy-Preserving Audit and Extraction of Digital Contents," Cryptology ePrint Archive, Report 2008/186, http://eprint.iacr.org, 2008.

[15]. G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Fourth Int'l Conf.Security and Privacy in Comm. Netowrks (SecureComm '08), pp. 1-10, 2008.

[16]. Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. 14th European Conf. Research in Computer

[17]. Security (ESORICS '09), pp. 355-370, 2009.

[18]. Swapnil V.Khedkar,A.D.Gawande, "Data Partitioning Technique to Improve Cloud

[19]. Data Storage Security." IJCSIT, Vol. 5 (3), 2014.

[20]. 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.

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

[22]. Sangram Rananavare, Prof. Anjali More, Pritam Vanne, Sneha Nanaware, "Enhanced Data Partitioning Technique for Improving CloudData Storage Security",International Journal of Modern Trends in Engineering and Research (IJMTER)Volume 02, Issue 01, [January - 2015].

[23]. Yogesh Shinde, Alka Vishwa, "Privacy Preserving using Data Partitioning Technique for Secure Cloud Storage",International Journal of Computer Applications (0975 – 8887) Volume 116 – No. 16, April 2015

[24]. Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011 .

[25]. Himika Parmar, Nancy Nainan and Sumaiya Thaseen, "Generation of secure one-time password based on image Authentication", Computer Science & Information Technology, pp. 195–206, 2012.

[26]. Huqun Wang, "Proxy Provable Data Possession in Public Clouds", IEEE Transactions On Services Computing, Vol. 6, No. 4, October-December 2013, ISSN: 1939-1374 .

[27]. Santosh Jogade, Ravi Sharma, Prof. Rajani Kadam, "Partitioning Data and Domain Integrity Checking for Storage - Improving Cloud Storage Security Using Data Partitioning Technique", International Journal of Emerging Research in Management &Technology, ISSN: 2278-9359 (Volume-3, Issue-3) .

[28]. 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 .

[29]. Kiran Gabhal1, Narendra Jadyal, Anurag More, Vinayak Bhalekar, 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.

[30]. C. Selvakumar, G. Jeeva Rathanam, M. R. Sumalatha, "PDDS - Improving Cloud Data Storage Security

[31]. Using Data Partitioning Technique", 2013 3rd IEEE International Advance Computing Conference (IACC)

[32]. Tiancheng Li; Ninghui Li; Jian Zhang; Molloy, I.; "Slicing: A NewApproach for Privacy Preserving Data Publishing," Knowledge and DataEngineering, IEEE Transactions on, vol.24, no.3, pp.561-574, March2012.

[33]. Paredes, L.N.G.; Zorzo, S.D.;, "Privacy Mechanism for Applications inCloud Computing," Latin America Transactions,

IEEE (Revista IEEEAmerica Latina) , vol.10, no.1, pp.1402-1407, Jan. 2012.

[34]. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving PublicAuditing for Storage Security in Cloud Computing," Proc. IEEEINFOCOM, Mar. 2010.

[35]. Nelson Gonzalez, Charles Miers, Fernando Red, Marcos," A quantitative analysis of current security concerns andsolutions for cloud computing", at Journal of Cloud Computing: Advances, Systems and Applications 2012.

[36]. Bhavna Makhija,VinitKumar gupta,Indrajit Rajput ,"Enhanced Data Security in Cloud Computing with Third PartyAuditor",Hasmukh Goswami College of Engineering,Vahelal,Gujarat, International Journal of Advanced Research inComputer Science and Software Engineering.

[37]. Ayad F. Barsoum and M. Anwar Hasan," Enabling Data Dynamic and Indirect Mutual Trust for Cloud ComputingStorage Systems", University of Waterloo, Ontario, Canada. IEEE TRANSACTIONS ON PARALLEL ANDDISTRIBUTED SYSYEMS VOL: PP NO: 99 YEAR 2013.

[38]. Sunil Sanka, Chittaranjan Hota, Muttukrishnan Rajarajan," Secure Data Access in Cloud Computing", ComputerScience and Information Systems Group, Birla Institute of Technology and Science-Pilani.