International Journal of Engineering & Technology, 7 (3.34) (2018) 905-907

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

An Enhanced Data Integrity Technique for Cloud Storage with Integrated Archive Using PDP

[1]C.Saravana Kumar, [2]K.VijayaKumar

[1]Associate Professor, Department of Information Technology St. Joseph's Institute of Technology, OMR, Chennai.
[2]Associate Professor, Department of Computer Science and Engineering St. Joseph's Institute of Technology, OMR, Chennai
**Corresponding Author E-mail: [1]mailtocsk@gmail.com, [2]mkvijay@msn.com*

Abstract

The cloud storage plays vital role in the recent trend in information technology. Storage in cloud to be efficient and effective with compressed files method but it faces problem which are related to space and time complexity. There are various techniques are available such as ZIP, JAR, and TXT etc with various disadvantages are exits due to incompatible for recovery records, Unicode names, and strong AES encryption. This problem has been addressed byusing some high compressibility ratio capabilities like RAR using WAR technology. It is more convenient of multipart archives, tight compression multimedia, text modes, strong AES encryption, recovery records to repair an archive in case of physical damage to file. It also includes all work that is related to web application like Servlet, XML, JSP, Image, HTML, CSS, and JSetc. The main advantage in WAR file is, it takes less time to transfer the file as it does all web related work at one time. Therefore RAR format using WAR file is used to have the storage of file in cloud which is of more credential than JAR format using ZIP file. The common security issues like confidentiality, integrity, Privacy-Provable Data Possession technique have been applied for better reliability.

Keywords:

1. Introduction

Cloud computing is a service oriented architecture paradigm which offer as a service manner. The cloud customer generates the request for accessing the resources at the cloud. These resources are offered by the cloud using resource allocation techniques through virtualization. There are various virtualization technique available bases on the demand and availability in the cloud computing data centre.The Virtual Machine Monitor (VM) performs the task for creating, terminating the scheduling the task. There are various domains are exists in the recent days which will gives the services to the society. The health care domain is a vital area in which the processing of collected data and store the data in the cloud. Cloud accessing needs the privacy and security with integrity.

The security has plays vital role in the field of cloud computing because the data are available at the remote end. The customer doesn't know about the data location and what level of security is available and applied over the customer data, so the integrity contribute a lot for the efficient cloud access. The integrity means the data cannot be changed by any other parties over the network.The healthcare data are maintained at the cloud server leads the performance problem due to latency occurs in the real world domain. This overcome using edge computing with cloud computing. The integrity of the cloud data is verified using Provable Data Possession (PDP) method with outsourcing support [1] [10].

The dynamic natures of data operations like update delete and insert operation are supported with scalability support with minor problem [2] [9]. This problem is addressed by using rank based data insertion feature with authenticated method with list of skipped form. It uses data structure in dynamic manner which suffers reply attacks and also indexed with block indexing mode [3] [8]. High-Availability and Integrity Layer (HAIL) method using error correction code with protection of integrity by using generic hash method of calculating hash value. It generated a Message Authentication Code (MAC) over the data in order to achieve strong protection from the mobile devices and networks. It provides the high success rate in availability with inter and intra domain access through redundancy [6] [4].

Multiple-Replica PDP (MR-PDP) technique has two approaches namely Single-Replica PDP and ENC-PDP method. The first method uses only one replica with collision attack in insecure manner. The second method prevents the collision over attack with security [5] [7]. The proposed system provides the data integrity through PDP in the form of compressed standard with WAR and JAR file combination. The data integrity related to the health care domain is addressed with reliability. The performance of the proposed system is improved by identifying various parameters in the medical field. The overall objective of the proposed system is to assess the security level of the cloud data in health care domain with the implementation of trust and integrity among the sensitive data.

The rest of the paper is organized as follows, Section 2 and section 3 represents the overall architecture and phases of the proposed algorithm. Section 4 and section 5 describes that the experimental evaluation and conclusion and future work.

2. Overall Architecture of the Proposed Model

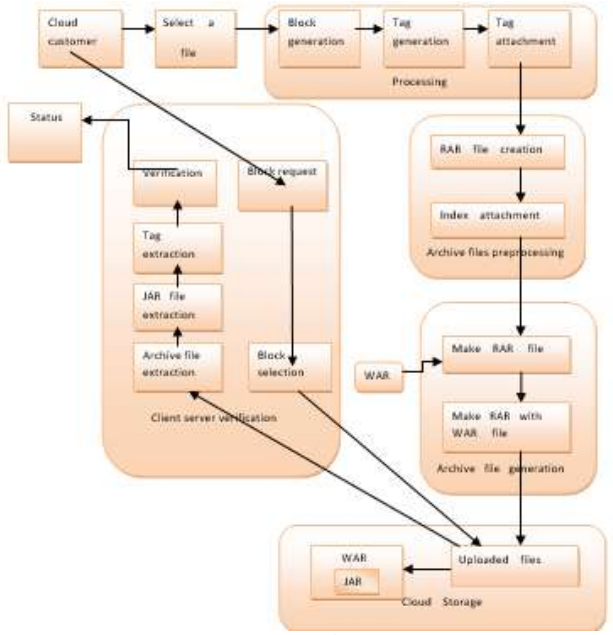

Figure 1: Architecture of the proposed model

Normally in the pre-process and store phase the client generates input file i.e., metadata which are normally stored in the client store and the client even modify the meta data and send them to the server and the modified data are stored in the server store.In the verify server possession phase the client generates the random challenge R and they are send to the server, where the server computes the proof of possession. If the proof of possession matching it sends the data back to the client and the client checks the correctness of the data that are said to be stored in the server store with the client Meta data. These files can be stored in the form of WAR and RAR with their corresponding tag generation and RAR format along with WAR which includes all work related to web applications like servlet, xml, jsp etc., The main advantage in WAR files are they take less time to transfer file and it does all web related works at one time.RAR along with WAR file used to have more credential than JAR format using ZIP files. It also includes PDP algorithms for generating public and secret key which are executed over the client side for processing the data. Secondly it has pre-process file of the meta-data. The random file treats input file using RAR and generate their corresponding tags for all blocks while processing.Thirdly generating the proof of RAR from the challenge to determine proof of possession.Finally Verifying proof from server whether the client verification of proof are valid or not. The overall architecture of the proposed model is shown in figure 1.

3. Phases of Proposed System

Preprocessing is the initial module. The customer's data will be stored in cloud from where the required file is selected for the challenge triggered by the data owner. These challenges searches for the block that were created in the form of compressed file (i.e.,) RAR in the cloud storage of data, those blocks were generated during the time of customer/client challenge. For each original file that are stored in cloud, there is a unique tag generated for the original files of metadata. During the process of call of challenge to the server, these unique tags for each Meta data of original file is actually attached at the initial stage itself.

After the processing step, the data's that are stored in cloud are actually of the file format of RAR. For each original file, the RAR file is created in order that it will be of high compressibility compared to other capabilities. There is a parameter called Index value which is uniquely set to all the WAR files such as RAR1.WAR,RAR2.WAR-in which 1,2,3,... are the unique index values set to the files. Here in archive file preprocessing module, index attachment also occurs along with the creation of RAR files.

Archive file generation mainly consist of making two possibilities of files such as RAR and WAR, where both RAR and WAR are brought into existence for use in cloud storage. The main motive for making RAR file is that it overcomes many disadvantages of ZIP file. As most of the files consist of the web content in it, here mainly WAR file is made that are used with RAR file format. Before combining RAR file with WAR, the data needed for WAR is actually outsourced from it. These are the two major generation of archive file.

The main purpose of verification is that to check whether the data that are stored in the cloud which belongs to the user is of high data integrity, security, confidentiality, privacy etc. This verification starts as soon as the customer/client sends the challenge to the server that request for blocks of data stored in cloud, then the particular block is selected from the cloud for uploaded files from the format of RAR using WAR, where WAR is nothing but the JAR files. Later these archive files are extracted which contains JAR file in it, initially we generated the tag and those tag attachment are actually verified to cross check with the Meta data of original file of the customer/client.

After the verification process is finished, the corresponding status is provided by the client/customer to the server to accept the verification process done on the server side in the cloud. If both the data's are matched correctly, the status provided to client will be indicated with the statement as true, if not with false and if any other problem other than matching it is indicated with null/zero statement during verification.

4. Experimental Evaluation

There are three customers have been interact over the cloud with different performance parameters such as source, destination, bandwidth and latency. Our evaluation we have considered three customers with various clustering of VM in order to utilize the VM effectively. Figure 2 shows that the configuration of the cloud for the proposed method.

ID	Amount	Image Size	Processin...	MIPS	RAM	Bandwidth	Priority
5	1	2000	2	950.0	1024	99800	1
4	1	1000	1	1000.0	512	100000	1
3	1	800	3	800.0	512	98100	2
2	1	400	3	1050.0	512	100200	3
1	2	900	4	1100.0	512	101200	4

Figure 2: Bandwidth Vs Latency

Figure 3 shows that the utilization of the resource with various fields with high reliability. The fields are VM ID, Cost of VM, Image size, MIPS, RAM size and current bandwidth and priority related the task andVM

Source	Destination	Bandwidth	Lantency
Customer3	Customer2	1.0	1.0
Customer3	Customer1	1.0	1.0
Customer3	VMClusteringt	1.0	1.0
Customer3	Datacenter1	1.0	1.0
Customer2	Customer3	1.0	1.0
Customer2	Customer1	1.0	1.0
Customer2	VMClusteringt	1.0	1.0
Customer2	Datacenter1	1.0	1.0
Customer1	Customer3	1.0	1.0
Customer1	Customer2	1.0	1.0
Customer1	VMClusteringt	1.0	1.0
Customer1	Datacenter1	1.0	1.0

Figure 3: various Fields

The VM energy consumption with health care domain is described as follows,

VM4-0 from **Host0** to **Host2** at **3.355** minutes. Source host was consuming 740000.0% of CPU, 204800.0% of RAM and 99.6% of power. Target host was consuming 0.0% of CPU, 0.0% of RAM and 0.0% of power.
VM5-0 from **Host0** to **Host1** at **19.654999999999998** minutes. Source host was consuming 200000.0% of CPU, 153600.0% of RAM and 78.0% of power. Target host was consuming 315000.0% of CPU, 51200.0% of RAM and 82.6% of power.
VM6-0 from **Host0** to **Host1** at **19.654999999999998** minutes. Source host was consuming 200000.0% of CPU, 153600.0% of RAM and 78.0% of power. Target host was consuming 315000.0% of CPU, 51200.0% of RAM and 82.6% of pow
VM4-4 from **Host0** to **Host2** at **19.654999999999998** minutes. Source host was consuming 200000.0% of CPU, 153600.0% of RAM and 78.0% of power. Target host was consuming 440000.0% of CPU, 51200.0% of RAM and 87.6% of power.
VM5-0 from **Host1** to **Host2** at **26.304333333333332** minutes. Source host was consuming 100000.0% of CPU, 51200.0% of RAM and 74.0% of power. Target host was consuming 100000.0% of CPU, 51200.0% of RAM and 74.0% of power.

5. Conclusion and Future Work

The cloud service has requested from the various user from different location using the cloud server as a data center. The cloud offers the service such as SaaS, PaaS, IaaS and XaaS and Software as a Service, Platform as a Service, Infrastructure as a Service and Everything as Service respectively. The proposed method of handling the cloud service in the health care domain provides the high quality services to the customer. The security also applied over the customer data but the user doesn't have knowledge about the security strength in the cloud which leads the performance problem. This problem also address by using the integrity mechanism applied over the data at the data center. The overall objective of the proposed method is to handle the customer data with more security and integrity by using PDP technique. In future the method has been extended to the big dataanalytics and machine learning related Artificial Intelligence system with high quality and accuracy.

References

[1] Ateniese, G.; Burns, R.; Curtmola, R.; Herring, J.; Kissner, L.; Peterson, Z. & Song, D., "Provable Data Possession at Untrusted Stores," in Proc. of the 14th ACM Conference on Computer and Communications Security, pp. 598—609, 2007.

[2] Ateniese, G.; Di Pietro, R.; Mancini, L. V. & Tsudik, G. (2008), "Scalable and Efficient Provable Data Possession", in Proc. of the 4th International Conference on Security and Privacy in Communication Network, pp. 9:1--9:10.

[3] Erway, C.; Küpçü, A.; Papamanthou, C. & Tamassia, R. (2009), "Dynamic Provable Data Possession," in Proc. of the 16th ACM Conference on Computer and Communications Security, ACM, New York, NY, USA, pp. 213—222.

[4] Bowers, K. D.; Juels, A. &Opera, A., "HAIL: A High-availability and Integrity Layer for Cloud Storage," in Proc. of the 16th ACM Conference on Computer and Communications Security, pp. 187—198, 2009

[5] Curtmola, R.; Khan, O.; Burns, R. & Ateniese, G., "MR-PDP: Multiple-Replica Provable Data Possession," in Proc. of Distributed Computing Systems, ICDCS '08. The 28th International Conference on', pp. 411-420, 2008

[6] Shacham, H. & Waters, B. (2008), "Compact Proofs of Retrievability," in Proc. of Josef Pieprzyk, ed., Advances in Cryptology - ASIACRYPT 2008, Springer Berlin Heidelberg, pp. 90-107.

[7] Ateniese, G.; Kamara, S. & Katz, J., "Proofs of Storage from Homomorphic Identification Protocols," in Proc. of the 15th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology', Springer-Verlag, Berlin, Heidelberg, pp. 319—333, 2009.

[8] Bowers, K. D.; Juels, A. & Oprea, A., "Proofs of Retrievability: Theory and Implementation," in Proc. of the 2009 ACM Workshop on Cloud Computing Security, pp. 43—54, 2009.

[9] Wang, Q.; Wang, C.; Li, J.; Ren, K. & Lou, W., "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," in Proc. of the 14th European Conference on Research in Computer Security, Springer-Verlag, Berlin, Heidelberg, pp. 355—370, 2009.

[10] Wang, C.; Wang, Q.; Ren, K. & Lou, W., "Privacy-preserving Public Auditing for Data Storage Security in Cloud Computing," in Proc. of the 29th Conference on Information Communications', IEEE Press, pp. 525—533, 2010.

[11] R.Joseph Manoj, M.D.Anto Praveena, K.Vijayakumar, "An ACO–ANN based feature selection algorithm for big data", Cluster Computing the Journal of Networks, Software Tools and Applications**,** ISSN: 1386-7857 (Print), 1573-7543 (Online) **DOI**: 10.1007/s10586-018-2550-z, 2018.

[12] K. Vijayakumar and C. Arun, "A Survey on Assessment of Risks in Cloud Migration", International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 10 No.66 May 2015.

[13] K. Vijayakumar, C.Arun,Automated risk identification using NLP in cloud based development environments Ambient Intell Human Computing, DOI 10.1007/s12652-017-0503-7, Springer May 2017.