Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

Effective Fragmentation Minimization by Cloud Enabled Back Up Storage

Dr. A. Pasumpon Pandian,
Professor, Computer Science Engineering,
KGiSL Institute of Technology, Coimbatore, India.
Email id: pasumponpandian32@gmail.com
Dr. S. Smys,
Professor, Department of Computer Science Engineering,
RVS technical Campus, Coimbatore, India.
Email id: smys375@gmail.com

Abstract: Nowadays to increase the efficiency, consistency and the quality of the organizations and to further extend the business world wide the digitization is followed in processing, storing and conveying the information. This in turn has also caused huge set of data flow paving way for the data recovery services. The cloud computing with the massive storage capabilities have become a predominantly used paradigm for data storage and recovery due to its on demand network access, elasticity, flexibility and pay as you go. Moreover to secure the information that is stored the information's are fragmented and stored. However this fragmentation process often occurs in the form of dispersed and scattered packages lacking proper order heightening the time and minimizing the efficiency of the recovery and information collection. To bring down the restoration time and enhance its efficiency the proposed method in the paper tries to reduce the fragmentation by minimizing the number of dispersed and scattered packages for this the paper utilizes the Hybridized Historical aware algorithm (HHAR) along with the cache aware filter to gather the historical information's associated with the back-up system and the identify the out of order containers respectively. Further the every data package is protected applying the advanced encryption standard by producing a key to authenticate the access of the data. The proposed model is simulated using the network simulator-II and the results obtained shows that the recovery time is enhanced by 95% and the restore performance is improved by 94.3%.

Keywords: Digitization, Data Fragmentation, Back up Storage, HHAR, Cache Aware Filter, AES Encryption.

Introduction

The world nowadays is facing a very high increase in the generation of the electronic data due to the digitization of the organization to improve the efficiency, consistency and the quality of the organization worldwide. This in turn demands an enormous volume of the storage devices to reserve a large quantity of data. The need for the storage facility leads to the emergence of the hard disk drive that can hold more data storing capabilities, but as the data generation exceeds the storage capacities of the hard disks, the user prefers vast storage facility and opts the cloud computing. The organizations usually owned separate storage of their own to back up their official details. Unfortunately most of the time the storage gets damaged or corrupted causing losses to the organization as the important information's of the organization are lost [1]. To avoid such situations the many research have brought in novel concepts following a plain back-up techniques, but this also was a failures due to the lack of convenience and reliability. Many new methods such as the REMUS [2] that provides "asynchronous virtual machine replication" HS-DRT [3] that combines fragmentation and duplication, Linux box [4] and SBBR [5] etc, where developed to provide a highly reliable and privacy protected back up. Despite their merits these methods were either too costly or

ISSN: 2582-337X (online)

Ubiquitous Computing & Communication Technologies

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

faced problems in maintenance and implementation. So to have a highly reliable and as well as secure back up storage the most organization prefers the cloud paradigm that offers service on pay per use.

The cloud computing has become a predominant, paradigm for rendering services according to the demands of the users; they offer several services such a platform, software, infrastructure, storage and recovery etc. Its cost effectiveness vast storage capability and computational capacity has made in popular among a wide range of applications and resources. The storage facilities of the cloud are incorporated with the duplication process to avoid the loss of the information even on the disaster that was natural or manmade. Further to protect the information stored in the cloud from unwanted misuse and hacking, it fragments the information followed by duplication to enhance the confidentiality and address the overloading issues of the cloud, Many a times these information's that are fragmented are not neatly arranged but randomly dispersed or scattered. This increase the recovery time of the information's as well as the efficiency in the restoration performance.

To handle the restoration performance issues and the recovery time of the information's that has occurred due to the scattered, dispersed and out of order information packages the proposed method utilizes the

(i) Historical aware rewriting integrated with the context based rewriting (HHAR (HAR +CBR)) to define, deduce and reduce the dispersed or the scattered containers,

(ii) Cache filtering (CAF) to acknowledge restored cache and

(iii) AES based encryption to authenticate the data usage.

The paper to minimize the fragmentation in the cloud storage and improve the restoring time and performance is organized with 2. The related works, 3. The Proposed method of data storage with recovery and minimized fragmentation, 4. The results evaluation, 5. Conclusion followed by reference.

Related Works

Sharma, Kruti et al [1] reviews the techniques "that are powerful solution in the form of online data backup and disaster recovery techniques and summarizes the techniques used in the cloud computing" Cully, et al [2] proposes an approach that includes a "software that is secured against the physical machine failures and performs parallel execution to simultaneously run the virtual machine that are active" Ueno, et al [3] the paper is about the "encryption evaluation in terms of the file data size and average response time, and further describing the prototype configuration of several practical applications" the figures.1 below shows the conventional back systems

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

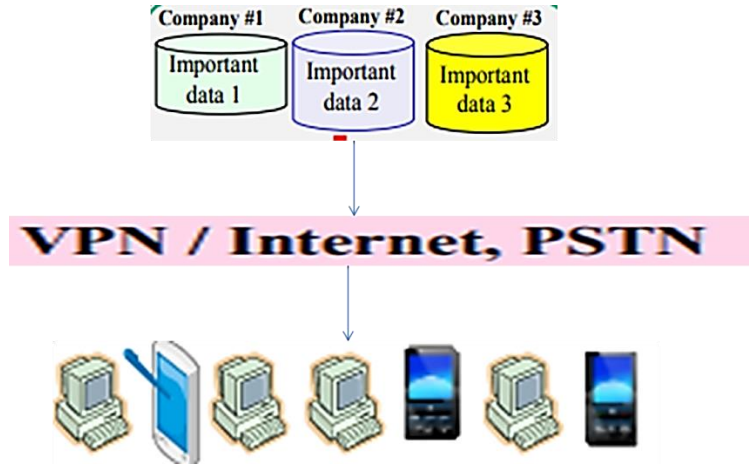

Figure.1 Conventional Back-Up System

Javaraiah, Vijaykumar et al [4] in his paper has introduced the online data back up along with the disaster recovery reducing the cost of the solution. Palkopoulou, Eleni et al [5] proposes an "analytical recovery time model based on GMPLS signaling and showed how the service imposed maximum outage on the configuration of the SBRR architecture" Kumar, et al [6] proposes the effective resource allocation process to enhance the performance competencies of the cloud. The proposed model of the paper takes the cryptographic ideas from the "Efficient cryptographic approach for data security in wireless sensor networks using MES VU." put forth by Praveena, A., and S. Smys et al [7] , the author Shi, Yong et al [8] presents the load balancing techniques by proposing the task scheduling strategy. Jennifer S. Raj et al [9] has proposed the model to manage the data flow in the adhoc network to improve its quality of service. Xia et al [10] is the study presenting the complete details of the past present and the future data duplication techniques and classifies the duplication into three types as location based, time based, level based Karthiban, K., et al [11] provides the details of the different privacy approaches to secure the data in the cloud computing. Han et al [12] presents the context aware distributed storage for the large amount of information's using the mobile cloud computing. Sridhar, S., et al [13] presents a hybrid level authentication scheme for the private cloud. Kaur et al [14] in his elaborates the details of the duplication techniques in the cloud to effectively manage its storing capabilities. Pandian, et al [15] provides the mobility management of the multiple networks in the IOT and the Sakthivel, et al [16] is the study on the necessity of fragmentation in the cloud back up storage. The security model for the data migration in the cloud was put forth by Shakya, Subarna et al [17] Chhabraa et al [18] is the study presenting the cloud storage data techniques with the duplication that is optimized.

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

Methodology

The proposed method utilizes the history aware rewriting [14] combined with the context based rewriting [18] to define and deduce and as well as reduce the scattered, out of order information containers and uses the cache filtering to acknowledge the restored cache and AES [19] based encryption to authenticate the data utilization. This section presents the methodology used in the proposed model to minimize the fragmentation. The figure.2 below shows the conventional cloud back-up

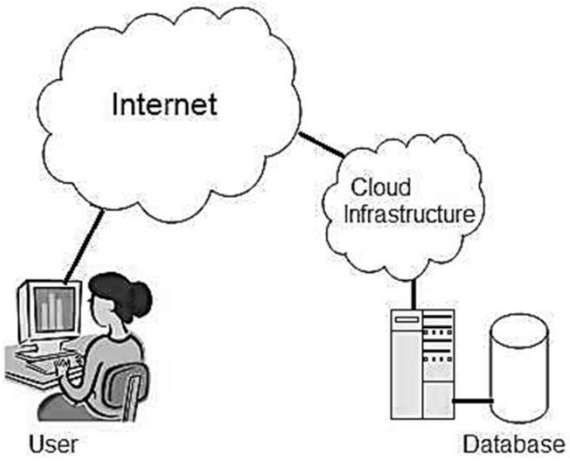

Figure .2 Conventional Cloud Storage [20]

Hybrid History Aware Rewriting and Authentication

The inherited frame work of the internal memory is structured with the identification numbers, for every information containers that is dispersed. This is loaded by the HHAR at the initial stage of the backup, to make sure the existence of the duplicate portions in theIh_s, are rewritten properly. This is achieved by the following the structure of the internal memory. The Em_s monitors the usage of every containers that are referred by the backup and records the usage of the data and its identification number and discards after the higher number of utilizations that are beyond the threshold values. Followed by the procedures of the HAR the CBR procedures are further applied to minimize the fragmentation by rewriting the fragmented replications. The CBR concentrates only on very highly fragmented replications utilizing two fixed-size contexts of the replicates, one is the disk context and the other is the stream context. Disk content is one block following the next and the stream context is the one stream following the next within a block. The intersection of the contexts makes fast the reading process because of the prefetching. The rewriting process tries to improve the similarity between the disk context and the stream context and utilizes the newly rewritten copy for the purpose of reading. The process initiates by setting the limits for the rewrite and then

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

makes rewrite decision based on the idea that the "credit of rewriting may be saved for another replication in the stream later or hoping the decision to rewrite would be done for a better replication in the future" i.e. the replication is done only for the highly fragmented replications and not for the replications within the rewrite limit. The block diagram below in figure.2 shows the phases in the proposed method to reduce the fragmented replicas.

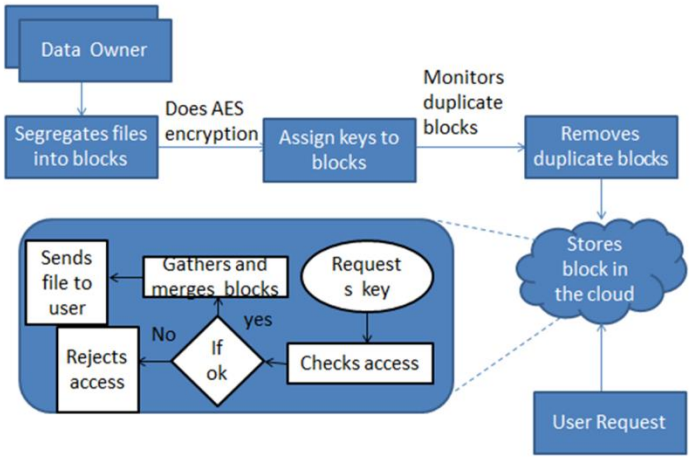

Figure.3 Block Diagram of Proposed

The above method does not care about the restore cache, so to acknowledge the restored cache the proffered method utilizes the cache aware filtering technique. The knowledge of the cache is utilized by the CAF to optimize the rewriting procedures. The CAF monitors the restores cache and stop the rewriting if the size of the information container is beyond the size of the cache used. For this the CAF estimates the restore cache size before making the decision on rewriting.

The contents to be duplicated are encrypted using the AES (advanced encryption standard) that is based on the substitution and the permutation of the network, where substitution and the permutation are the mathematical functions carried out in the block ciphers. The figure.4 below show the fundamental structure of the advanced encryption standard, the secret key generation for the text is done using this procedure and sent to the client , and this secret key acts an authentication to the access of the data and is used only by the authorized persons.

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

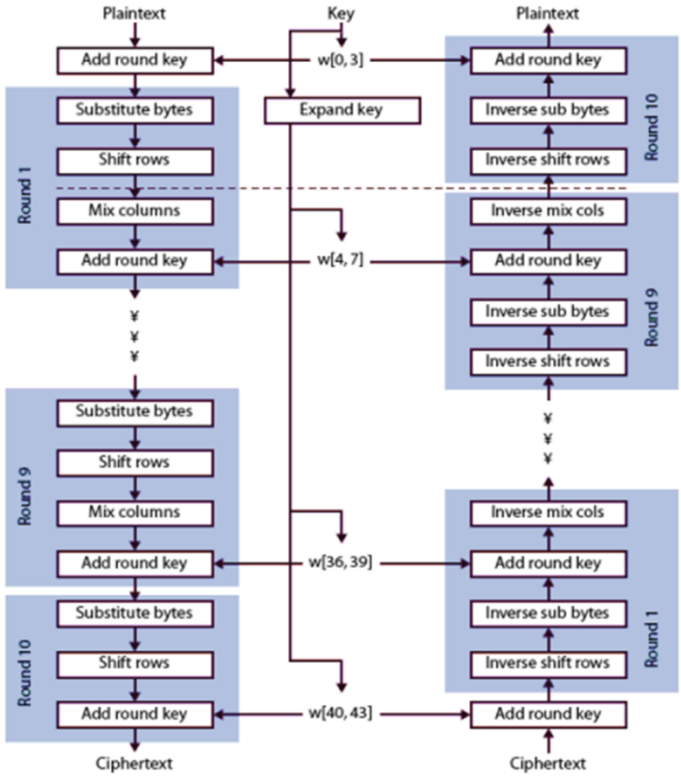

Figure.4 Fundamental Structure of the AES

Simulation Results

The proposed frame work developed using the HHAR (HAR+ CBR +CAF) minimizes the fragmentation by the process of rewriting and properly deciding the blocks that are to be rewritten based on the cache availability thus improving the storage capacity, bandwidth utilization , server space utilization along with the enhancements in the restore performance and the storage efficiency. This minimizes the total count that are blocks that are be stored in the cloud decreasing the write operation and time required for the restoration. The proposed model is simulated using the network simulator-II in terms of storage capacity, server space utilization; restore performance and storage efficiency acquired for different number of blocks that holds information of different data sets.

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

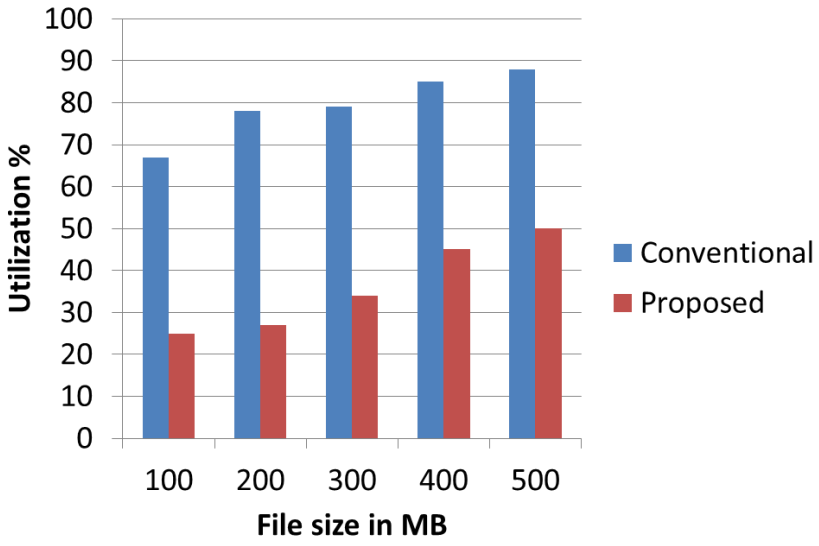

Figure.5 Server Space Utilization

The figure.5 represents the simulation results of the server space utilization for the conventional and the proposed method. The proposed method shows better server utilization due to the reduced number of blocks stored to the cloud by storing only a single instance of blocks before the duplication. The secret key generation for each block enables to identify the repeated blocks there reducing the server utilization.

File sizes	Conventional			Proposed		
	Restore Performance %	Storage efficiency %	Authentication percentage %	Restore Performance %	Storage efficiency %	Authentication percentage %
100	20	23	24	76	75	94
200	12	25	25	78	80	90
300	15	27.6	34	80	85	89
400	17	28.5	33	85	89.4	85
500	19	35	31	84	90	84

Table .1 Performance Comparisons

ISSN: 2582-337X (online)

Journal of Ubiquitous Computing and Communication Technologies (UCCT) (2020)
Vol.02/ No. 01
Pages: 1- 9
https://www.irojournals.com/jucct/
DOI: https://doi.org/10.36548/jucct.2020.1.001

The table.1 is the compares the restore performance, the storage efficiency and the authentication percentage of the proposed model and the conventional system. The results show that the proposed model is performs better, enhancing the restore performance, the storage efficiency and the authentication percentage than the normal cloud back-up.

Conclusion

As the back-up of information's and the duplication of the data are very much necessitated in the today's world, the cloud back up is mostly preferred but to enhance the security of the information stored in cloud the fragmentation is done causing much increase in the restoration time. So to minimize the restoration time and improve the storage efficiency, the propose model puts forward the effective fragmentation minimization techniques utilizing the HHAR and the AES , the results observed showed that the proposed method provided a better restoration time and improve the storage efficiency, compare do the conventional back up.

References

[1] Sharma, Kruti, and Kavita R. Singh. "Online data back-up and disaster recovery techniques in cloud computing: A review." *International Journal of Engineering and Innovative Technology (IJEIT)* 2, no. 5 (2012): 249-254.

[2] Cully, Brendan, Geoffrey Lefebvre, Dutch Meyer, Mike Feeley, Norm Hutchinson, and Andrew Warfield. "Remus: High availability via asynchronous virtual machine replication." In *Proceedings of the 5th USENIX symposium on networked systems design and implementation*, pp. 161-174. 2008.

[3] Ueno, Yoichiro, Noriharu Miyaho, Shuichi Suzuki, and Kazuo Ichihara. "Performance evaluation of a disaster recovery system and practical network system applications." In *2010 Fifth International Conference on Systems and Networks Communications*, pp. 195-200. IEEE, 2010.

[4] Javaraiah, Vijaykumar. "Backup for cloud and disaster recovery for consumers and SMBs." In *2011 Fifth IEEE International Conference on Advanced Telecommunication Systems and Networks (ANTS)*, pp. 1-3. IEEE, 2011.

[5] Palkopoulou, Eleni, Dominic A. Schupke, and Thomas Bauschert. "Recovery Time Analysis for the Shared Backup Router Resources (SBRR) Architecture." In *2011 IEEE International Conference on Communications (ICC)*, pp. 1-6. IEEE, 2011.

[6] Kumar, T. Senthil. "Efficient resource allocation and QOS enhancements of IoT with FOG network." *J ISMAC* 1 (2019): 101-110.

[7] Praveena, A., and S. Smys. "Efficient cryptographic approach for data security in wireless sensor networks using MES VU." In *2016 10th international conference on intelligent systems and control (ISCO)*, pp. 1-6. IEEE, 2016.

ISSN: 2582-337X (online)

[8] Shi, Yong, and Kai Qian. "LBMM: a load balancing based task scheduling algorithm for cloud." In *Future of Information and Communication Conference*, pp. 706-712. Springer, Cham, 2019.

[9] Smys, S., and Jennifer S. Raj. "A STOCHASTIC MOBILE DATA TRAFFIC MODEL FOR VEHICULAR AD HOC NETWORKS." *Journal of Ubiquitous Computing and Communication Technologies (UCCT)* 1, no. 01 (2019): 55-63.

[10] Xia, Wen, Hong Jiang, Dan Feng, Fred Douglis, Philip Shilane, Yu Hua, Min Fu, Yucheng Zhang, and Yukun Zhou. "A comprehensive study of the past, present, and future of data deduplication." *Proceedings of the IEEE* 104, no. 9 (2016): 1681-1710.

[11] Karthiban, K., and S. Smys. "Privacy preserving approaches in cloud computing." In *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 462-467. IEEE, 2018.

[12] Han, Dong, Ye Yan, and Tao Shu. "Context-aware distributed storage in mobile cloud computing." In *2015 IEEE Globecom Workshops (GC Wkshps)*, pp. 1-6. IEEE, 2015.

[13] Sridhar, S., and S. Smys. "A hybrid multilevel authentication scheme for private cloud environment." In *2016 10th International Conference on Intelligent Systems and Control (ISCO)*, pp. 1-5. IEEE, 2016.

[14] Kaur, Ravneet, Inderveer Chana, and Jhilik Bhattacharya. "Data deduplication techniques for efficient cloud storage management: a systematic review." *The Journal of Supercomputing* 74, no. 5 (2018): 2035-2085.

[15] Pandian, M. Durai. "ENHANCED NETWORK PERFORMANCE AND MOBILITY MANAGEMENT OF IOT MULTI NETWORKS." *Journal of trends in Computer Science and Smart technology (TCSST)* 1, no. 02 (2019): 95-105.

[16] Sakthivel, M., Karnajoy Santal, and K. Bhaskar Rao. "Fragmentation Study for Deduplication in cache Backup Storage."

[17] Shakya, Subarna. "AN EFFICIENT SECURITY FRAMEWORK FOR DATA MIGRATION IN A CLOUD COMPUTING ENVIRONMENT." *Journal of Artificial Intelligence* 1, no. 01 (2019): 45-53.

[18] Chhabraa, Nipun, and Manju Balab. "Study of existing cloud data storage Techniques with respect to optimized Duplication: Deduplication." (2019).

[19] Abdullah, Ako Muhamad. "Advanced encryption standard (AES) algorithm to encrypt and decrypt data." *Cryptography and Network Security* 16 (2017).

[20] https://www.rfwireless-world.com/Tutorials/traditional-storage-vs-cloud-storage.html

ISSN: 2582-337X (online)

Ubiquitous Computing & Communication Technologies