EasyChair Preprint

№ 3313

Breaking down and Reduplication of Information in Cloud for Best Overall Performance and Protection

Nageswara Rao Gali, Venkateswarlu Bondu and
Rao G. Jagadeeswararao

May 2, 2020

Breaking down and Reduplication of information in Cloud for best overall performance and protection

Gali Nageswara Rao[1], Bondu Venkateswarlu[2], JagadeeswaraRao.G[3]
[1,3]Department of Information Technology, Aditya Institute of Technology and Management, Tekkali
[2]Department of Computer Engineering, Dayananda Sagar University, Bengaluru, India.

Abstract

As an increase in the usage of the database, the data security and storage of data becomes a very big issue. To overcome this, cloud computing comes first. All the data will be stored in a third party location and retrieve whenever the user wants to access it. In order to achieve this, we came up with "Disintegration and reduplication of data in cloud for safety and security". In this methodology, when the user sends the file on the cloud server it gets fragmented. Fragmentation is a process of dividing the file into some fragments in a way that it is impossible to attack the total file at a time. Each node stores a single fragment of a particular file. The main aim of this work is to give security, protection, and performance against all types of attacks.

Keywords: *Fragmentation, Reduplication, Cloud security, Encryption, decryption, performance.*

1. Introduction

Traditional business applications are expensive and complicated. The amount of hardware and software required to run them are difficult. With cloud computing, you can eliminate the problems that come with storing of data because you are not managing hardware and software that becomes the responsibility of the software vendor. Cloud computing is a platform that is Configuring, manipulating, and accessing the hardware and software resources remotely. It offers online infrastructure, data storage, and software applications. In cloud computing, the software is not going to be installed locally on the private computer, i.e., it is platform-independent. Cloud can have four sorts of access: Public, private, hybrid and community. Public cloud is that in which entire infrastructure is found on the premises of a cloud computing company that gives the cloud surface. Due to this the safety level is low. In a private cloud, all the computing infrastructure is with yourself and not shared. Due to this the safety and control level is high. In a hybrid cloud, it comprises both public and personal cloud, user can use depending upon their purpose. In a community cloud, the cloud is shared between an organization that matches into a selected community like geographic community, professional community, etc. The cloud services were categorized into majorly three types: IaaS, PaaS, and SaaS.

The IT infrastructure usage and maintenance is completely changed after development of these cloud services [1]. Cloud computing had become a popular computing model, which provides on-demand dynamic service by facilitating virtual hardware and software resources [2]. The data that was stored in cloud severs must be secured and allow for modification based the authenticity of the users. The cloud was built upon basic cryptographic technique, which will encrypts any data files coming into the cloud to guarantee their privacy and security [3]. Cloud computing provides a large amount of storage for different clients. The main advantages are: Users can access, manipulate, and configure the applications as utilities over the internet at any time. It is more reliable as it offers a load balance. Cloud provides low cost and good reliability [4, 16].

To develop effective and efficient end solutions first we need to understand the security and privacy risks in the cloud. All though cloud allows its users to avoid initial costs, less operating costs, and increase their activity by immediately acquiring infrastructural resources, services when needed, their unique architectural features raise many security and privacy concerns [5]. Whether it is a public, private, or hybrid cloud, cloud security is the protection of data involved in cloud computing. The security measures protect customers' privacy as well as framing authentication rules for individual users and other

devices.

Cryptography is often mentioned because of the practice of the study of hiding and securing the data. Cryptography is the science of keeping data secret and safe. There has always been an urge among humans to stay sensitive data safe and secure in order that it might not cause unwanted intrusion into sensitive information which could lead to severe problems. Thus, cryptography has been practiced by humans from olden times to keep their data secure [6].

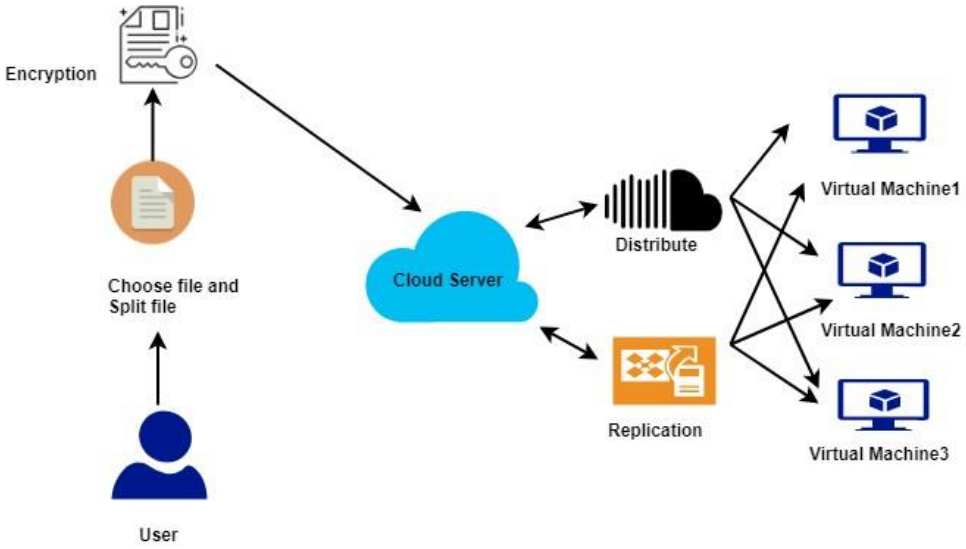

Fig.1 Methodology

The Cloud computing paradigm allows accessing the services and data from remote data centers. Now a days companies are moving data to the cloud, and they are trying to protect their data against threats.

The major security goals are confidentiality, integrity, and availability. Cryptography is a mixture of three kinds of algorithms: (1) Symmetric Key Algorithms (2) Asymmetric Key Algorithms and (3) hashing. Data cryptography is nothing but encoding the content of the information like text and media from understandable, meaningless, and invisible during this data transfer and storage, this conversion of information from understandable to not understandable format is referred to as encryption. Decipherment is the process of getting the initial information from Encrypted information. To encipher information on cloud storage each symmetric key and asymmetric key are often used, but consistent with the large size of the database and data stored in cloud storage using the symmetric-key algorithm is quicker than asymmetric key [7].

The information in the Cloud will partitioned for Security and those partitions are commonly called fragments, and these fragments will be allocated at strategic locations among the cloud data servers. The fragmentation of a file will be done in a way that, no individual fragments may not contain any meaningful data and this will be performed based upon the user criteria. A cloud node may be anything which can be a computing VM, storage VM. These nodes contains a definite fragment to extend the information security. An attack on one node shouldn't reveal the locations of various fragments in the cloud. We have to keep our file fragments at different locations and unaware from an attacker. The selection of the cloud nodes are also important I this context and they are not neighbors and they are at a meaningful distance from each other. [8].

The community should take predefined measures to ensure security. A displacement of the file information exists to adopt global standards (for example, open-source) to make sure interoperability among service providers. Even if the community at massive is aware of the necessity for security and is trying to initiate strong measures [9].

Storing data in an exceedingly public cloud could be one more security issue in cloud computing. Normally clouds will be implemented at centralized storage facilities like data centers, which may be a selective target for hackers. Public clouds are more

vulnerable for the attacks. Maintaining a personal cloud if possible for very sensitive data is recommended to avoid these security issues. [10].

Privacy and security are the main concerns in the usage of the cloud for data. Features of knowledge like privacy, integrity, and protection must be upheld. Different mechanisms and policies are employed by different service providers counting on the sort of knowledge, size of knowledge, and nature of knowledge. In Cloud Computing, the most advantage is that various organizations can share the info. But, the advantage causes data to be in danger. Therefore, the risk to the info must be overcome and thus data protection is extremely much required [11]. Our main contributions in this paper are as allows: we propose Disintegration and Reduplication of Data in Cloud for Optimal Performance and Security that collectively approaches the safety and performance issues.

First the file is going to be fragmented and these file fragments will be distributed over the different cloud nodes. To ensure the security of the file, each cloud node stores singly one fragment of a specific record, this makes no data has been revealed to the attacker.

1.1 Objective of the Paper

- To achieve optimal performance and security by division and replication of data in the cloud.
- We break the file into several fragments and replicate the fragments over the cloud nodes.

The sections in this paper are organized as follows. In Section 2 we gave an overview of the related work. In Section 3, the preliminary concepts of the work. Section 4 discusses the System model and methodology. Section 5 gives the experimental setup and results, and Section 6 is concluding the paper.

2. Related work

Juels et al. [12] introduced an approach of Iris document framework for the information movement to the cloud is performed by the Iris document framework. A passage application is structured and utilized inside the association that guarantees the honesty and freshness of the information utilizing a Merkle tree. The document squares, MAC codes, and form numbers are put away at different degrees of the tree. The proposed procedure in [12] vigorously relies upon the user′ s utilized to plan for information secrecy. Additionally, the likely measure of misfortune just if there should arise an occurrence of information hardening because of interruption or access by different VMs can't be diminished. Our proposed methodology doesn't rely on ordinary cryptographic systems for information security. In addition, the proposed approach doesn't store the whole record on one hub to stay away from the bargain of the entirety of the data just in the event of a fruitful assault on the hub. The creators in [13] drew nearer the multi-tenure related issues in the distributed storage by using the combined stockpiling and local access control. The Dike approval configuration is arranged that blends the local access to the executives and subsequently the inhabitant name house separation. The arranged framework is assumed and works for object-based document frameworks. Nonetheless, the spillage of basic data just if there should arise an occurrence of ill-advised disinfection and malignant VM isn't taken care of. The proposed technique handles the spillage of basic data by dividing documents and utilizing different hubs to store one record.

Wayne A. Jansen [14] proposed the method where the issues of the cloud are sorted out into a few general classes: trust, engineering, personality the board, Software disengagement, information insurance, and accessibility. Episodes may include different sorts of extortion, the harm of information assets, and robbery of information by present or previous representatives, temporary workers, and different gatherings that have gotten access to other private information. It is not possible to confirm the correct working of a subsystem and in this manner, the viability of security controls as broadly as a hierarchical framework. Both the customer and server-side security is unavoidable in distributed storage. It is adequate to expand the association's distinguishing proof and verification structure as information affectability, and security of information has progressively become a need for associations, and unapproved access to data assets

inside the cloud could likewise be a difficult issue.

Information segregation and information area are significant inside the information security part inside the cloud. The data put away on the cloud hubs ought to be accessible for approved clients at any expense whenever. Yang Tang, et al. [15] proposed a downsize information the board costs, this paper depicts redistributing information reinforcements off-site to outsider distributed storage administrations. Blur is to accomplish fine-grained, strategy based access control, and document guaranteed erasure. Blur connects the re-appropriated records with document get to approaches and it erases documents to make them hopeless to heaps of record get to strategies.

3. Preliminary Concepts

Before we dive into the proposed methodology, let's look at the related concepts in the following.

3.1 Data Encryption and Decryption

Encryption is the process of converting plain text information into a not understandable format meaningless data called ciphertext. Decryption is the procedure of converting ciphertext again to plaintext. To encrypt more than a small number of facts, symmetric encryption is used. The symmetric key is used at some stage in each of the encryption and decryption processes. The intention of each encryption is to make it as difficult as feasible to decrypt the generated ciphertext. The longer the key, the harder it is to decrypt a piece of ciphertext.

3.2 Elliptic Curve Cryptography (ECC)

One of the approaches to public-key cryptography is Elliptic Curve Cryptography. This technique was first proposed individually by Neal Koblitz and Victor Miller in 1985. The ECC is based on the Elliptic Curve Discrete Logarithm problem, which is a known NP-Hard problem. An elliptic curve is defined by the equation,

$$y^2 = x^3 + ax + b$$

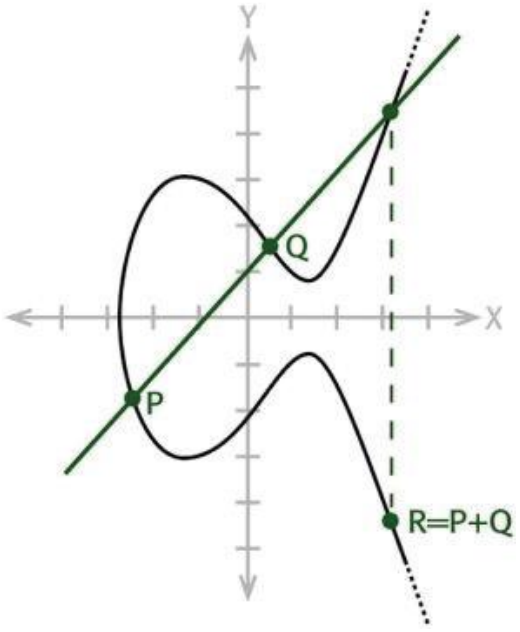

Fig 2: Elliptic curve showing the operation P + Q = R.

3.2.1 Basic Algorithm

Alice, Bob agree on a (non-secret) elliptic curve and a (non-secret) fixed curve point F. Alice chooses a secret random integer A_k which is her secret key, and publishes the curve point $A_P = A_kF$ as her public key. Bob will do the same, let $B_p = B_kF$. Consider a scenario, Alice need to communicate with Bob. One way is, Alice has to simply compute, $R_k = A_kB_P$. The result R_k will be used as the secret key for a conventional symmetric block cipher (say DES). Bob can compute the same number by calculating B_kA_P [17],

Since $B_kA_P = B_k \cdot (A_kF) = A_k \cdot (B_kF) = A_kB_P$.

The security lies in this algorithm is that, it is difficult to compute k for a given F and kF.

3.2.2 RSA vs ECC

Menezes and Jurisic [18], mentioned following observations by which we can assess the working of RSA and ECC. To get a considerable amount of security, 160-bit modulus should be sufficient for ECC, whereas RSA system takes 1024-bit modulus.

Time to break (in MIPS–Years)	RSA key-size (in bits)	ECC key-size (in bits)
10^4	512	106
10^8	768	132
10^{11}	1024	160
10^{20}	2048	210
10^{78}	21000	600

Table 1: Comparison of strength of RSA and ECC

4. System Model

4.1 System architecture

System architecture mainly consists of the user, cloud server, and server machines.

USER:
1) User makes registration by providing their details.
2) Users will log in by using a username and password.
3) After login, they choose the file and split that file into the number of fragments.
4) After splitting the files into fragments, they can view the fragments and details of the fragments.
5) Next, distribute the fragments by selecting distribute operation to the server machines.
SERVER:
1) Receive the fragments from the user which is converted into the encrypted format and store the fragments. The server receives the request details from the user and gives a response for the particular request what the user has requested.
2) Disintegration and Reduplication of Data in the Cloud can be done by the servers.

The algorithm 1 is for data encryption of the given file which takes a random number and multiplies it with a prime number which will be stored in a big integer format. Here we are using the ECC algorithm for encryption. Algorithm 2 decrypts the encrypted messages where the ciphertext is converted into a string.

4.2 Algorithm1 : Algorithm for Data Encryption

```
Random r1 = new Random();
BigInteger I = BigInteger.probablePrime(3, r1);
BigInteger J = BigInteger.probablePrime(3, r1);
BigInteger K =I.multiply(J);
Random rand = new Random();
int kval= K.intValue();
int kres=rand.nextInt(kval-1);
BigInteger  p=BigInteger.valueOf(kres);
BigInteger i1=k.multiply(I);
BigInteger N=new BigInteger(msg.getBytes());
BigInteger i2=N.add(i1);
encmsg=i1+","+i2;
```

Algorithm2 : Algorithm for Data Decryption

```
String spt[]=cipher.split(",");
BigInteger i1=new BigInteger(spt[0]);
BigInteger i2=new BigInteger(spt[1]);
BigInteger n=i2.subtract (i1);
String filedata=new String (n.toByteArray());
return filedata;
```

5. Result Analysis

5.1 File content size v/s No. of fragments

File content size	No .of Fragments
275	1
854	1
5482	6
7543	8
11642	12

Table 2: File content size v/s No. of fragments

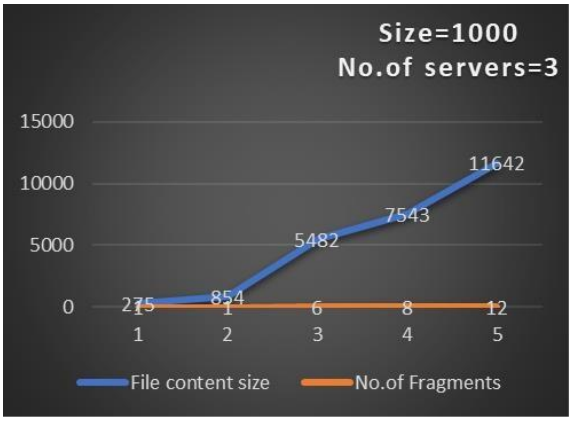

Fig 3: File content size v/s No. of fragments

5.2 For Performance

5.2.1 Read, Replicate and Retrieval time comparison

No. of fragments	Read and replicate time for our methodology(ms)	Read and replicate time for DROPs methodology(ms)
1	83	80
6	63	76
8	69	78
12	78	83

Table 3. Read and replicate time comparison

No. of fragments	Retrieval time for our methodology(ms)	Retrieval time for DROPs
1	46	32
6	16	27
8	24	30

12	47	31

Table 4. Retrieval time comparison

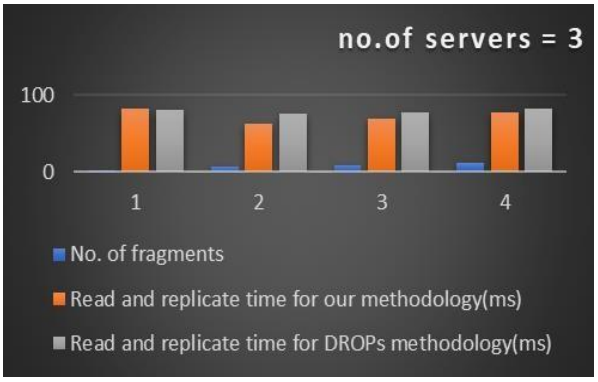

Fig 4: File content size v/s No. of fragments

Here in graph Fig.4 we see the read and replicate time for our methodology and drops methodology and we see the read and replicate time is optimal when the fragments are between 6-8 and therefore gives better performance than drops for those numbers of fragments.

5.2.2 Retrieval time comparison

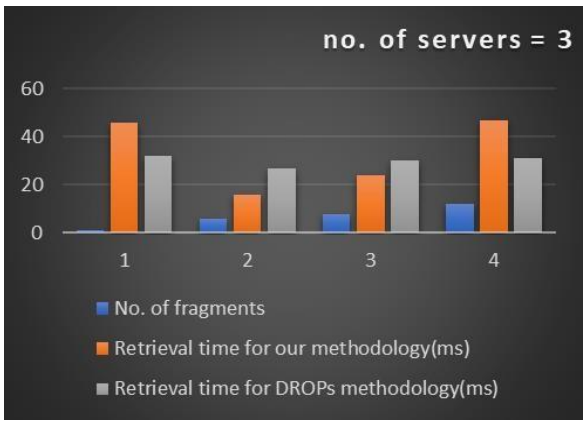

Fig 5: comparison with DROPS

Here in graph Fig.5 we see the retrieval time for our methodology and drops methodology and we see the retrieval time is optimal when the fragments are between 6-8 and therefore gives better performance than drops for those number of fragments. But there is a slight drop in the overall performance.

6. Conclusion

In this paper, we proposed a model for high security through file disintegration, encryption, and performance by retrieval time. This is achieved by dividing the file into small chunks and distributing it in different nodes of the cloud. The data inside the fragments will be converted into an encrypted format for better security and also contains one replication of that data, so that if there is any sort of attack no meaningful information is revealed, only corrupted information is gained. As shown in the result the comparison between our model and drops technology gives detailed information about performance through minimizing retrieval time for some range of fragments. Therefore, the obtained results increase the security of the user files, avoids the success of unauthorized party attacks of accessing the data, also in better optimizing the performance. User can download their file successfully in the original format after decrypting the data inside the

file.

References

[1] DROPS: Division and Replication of Data in Cloud for Optimal Performance and Security-Mazhar Ali, Student Member, IEEE, Kashif Bilal, Student Member, IEEE, Samee U. Khan, SeniorMember, IEEE, BharadwajVeeravalli, Senior Member, IEEE, Keqin Li, Senior Member, IEEE, and Albert Y. Zomaya, Fellow, IEEE Journals & Magazines (Volume: 6, Issue: 2, Year:2018)

[2] Addressing cloud computing security issues - Dimitrios Zissis DimitriosLekkas Publishedin:Future Generation Computer Systems archive Volume 28 Issue 3, March, 2012, Pages 583-592.

[3] Y. Tang, P. P. Lee, J. C. S. Lui, and R. Perlman, "Secure overlay cloud storage with access control and assured deletion," IEEE Transactions on Dependable and Secure Computing, Vol. 9, No. 6,Nov. 2012, pp. 903-916.

[4] A. N. Khan, M. L. M. Kiah, S. U. Khan, and S. A. Madani,"Towards Secure Mobile Cloud Computing: A Survey," Future Generation Computer Systems, Vol. 29, No. 5, 2013, pp.1278-1299.

[5] Security and Privacy Challenges in Cloud Computing Environments-Hassan Takabi,JamesB. D. Joshi,Article in IEEE Security and Privacy Magazine · January2011.

[6] Parul Agarwal Volume 8, No. 5, May-June 2017 International Journal of Advanced Research in Computer Science RESEARCH PAPER Available Online at www.ijarcs.info © 2015-19, IJARCS All Rights Reserved 2193 ISSN No. 0976-5697 Cryptography Based Security for Cloud Computing System

[7] Using Cryptography Algorithms to Secure Cloud Computing Data and Services Eng. Hashem H. Ramadan, Moussa Adamou Djamilou. American Journal of Engineering Research (AJER) e-ISSN: 2320-0847 p-ISSN : 2320-0936 Volume-6, Issue-10, pp-334-337

[8] A. N. Khan, M.L. M. Kiah, S. A. Madani, and M. Ali, "Enhanced dynamic credential generation scheme for protection of user identity in mobile-cloud computing, The Journal of Supercomputing, Vol. 66, No. 3, 2013,pp. 1687-1706.

[9] L. M. Kaufman, "Data security in the world of cloud computing," IEEE Security and Privacy, Vol. 7, No. 4, 2009, pp.61-64

[10] Data Security in Cloud Computing, Ahmed AlbugmiMadini O. Alassafi Robert Walters, Gary Wills.

[11] Protection and Security of Data in Cloud Computing, E. Poonguzhali 2. Suhas Rao M V, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by, www.ijert.org ICPCN - 2017 ConferenceProceedings.

[12] A. Juels and A. Opera, "New approaches to security and availability for cloud data," Communications ofthe ACM, Vol. 56, No. 2, 2013, pp.64-73.

[13] G. Kappes, A. Hatzieleftheriou, and S. V. Anastasiadis, "Dike: Virtualization-aware Access Control for Multitenant Filesystems," University of Ioannina, Greece, Technical Report No. DCS2013-1, 2013

[14] Wayne A. Jansen, "Cloud Hooks: Security and Privacy Issues in Cloud Computing" Proceedings of the 44th Hawaii International Conference on System Sciences -2011.

[15] Yang Tang, Patrick P.C. Lee, John C.S.Lui, Radia Perlman, "Secure Overlay Cloud Storage with Access Control and Assured Deletion" IEEE transactions on dependable and secure computing, vol.9, No.6, November/December 2012.

[16] G. J. Rao and G. S. Babu, "Energy analysis of task scheduling algorithms in green cloud," 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, 2017, pp. 302-305.

[17] Vivek Kapoor, Vivek Sonny Abraham, Ramesh Singh, "Elliptic Curve Cryptography", ACM Ubiquity, Volume 9, Issue 20, May 20 – 26, 2008

[18] Aleksandar Jurisic and Alfred J. Menezes. Elliptic curves and cryptography. Dr. Dobb's Journal, 1997.