ISSN: 2321-7782 (Online)
Volume 3, Issue 12, December 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Cumulative Study for Three Servers Swapping in Cloud

Ashwini Raosaheb Taksal[1]
Dept. Of Computer Science and Engineering
Bhivarabai Sawant Institute of Technology & Research
Pune, India

Sonali A. Patil[2]
Dept. Of Computer Science and Engineering
Bhivarabai Sawant Institute of Technology & Research
Pune, India

Abstract: ***Cloud computing plays vital role in today's IT industry, as it offers tremendous computing and storage facilities for the task and outsourced data respectively. To provide these facilities cloud is powered with complex integration of huge number of servers, due to this cloud computing becomes gigantic environment for computing which eventually makes it as one of the most important driving force for the computer industry. Due to huge number of the servers in cloud computing it compromises with the poor security for the stored data, but applying of cryptographic techniques solve this to the great extent. Even though the threat of data accessing is always there, this is directly depend on the storage pattern of the data in different servers of cloud. So it is necessary to wipeout traces of data expedition to the destination server to avoid accessing through any illegal means. As solution to this data swapping between server draws more attention as it successfully wipeout the traces. So a need of the better system is always there which enforces re-allocation of data by constant swapping between the many servers to provide un-predictable data route for the illegal seekers, which shall make impossible to access the cloud confidential data in any illegal means.***

Keywords: content confidentiality, access confidentiality, pattern confidentiality, distributed swapping.

I. INTRODUCTION

In today's era of it revolution size of the data being used for different purpose is increasing tremendously. With this increasing number the overhead to maintain such data is also increasing. So to reduce down such overhead, companies started relying on cloud platforms. As cloud offers number of advantages in data storing, data accessing with proper data management.

Also the end users have the best experience with the cloud as they can access the data anytime anywhere even on their smartphones. Very good examples of such services are amazon cloud, Google app engine and azure of Microsoft. Page Layout Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

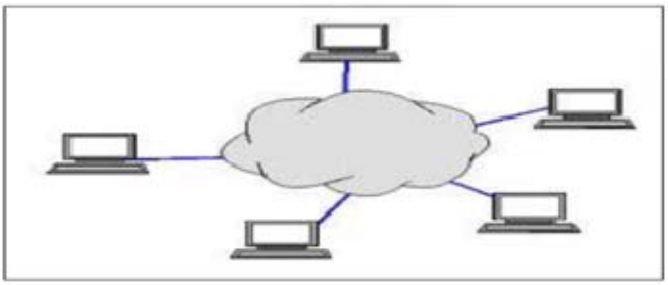

Figure 1: Simple Cloud Architecture

Such application offers storage for users at very low cost. E.g. the average costing of amazon is $0.16 per gigabyte-month, which is relatively low compared to the cost required for creating such storing infrastructures therefore more and more users are adopted it as an option for storing the data.

Several techniques were being proposed to maintain access confidentiality. The common thing between all those methods is to cut the link between the data and the location where data is being stored. By taking the advantages provided by cloud service providers, number of risks also coming. One of the major risks is the confidentiality of the cloud data. Cloud data confidentiality is getting affected due to various reasons such as software bottlenecks, operator's bugs, and external vulnerabilities. From those entire proposed approach shuffle index is the one which got lots of attention. Shuffle index provides a hierarchical manipulation of data based on keys. While doing this it preserves the access confidentiality of the data.

Access Confidentiality

Access confidentiality is term given to the ability of giving specific access rights to the specific node or data in cloud system against the server. Data confidentiality mainly comes in pictures when the entities outsourced the data for storage on cloud. There are many applications for which this issue is not only affects their privacy but also there juristic concerns. So it becomes crucial to deal with such issue. Index structure of static encrypted form does not possess the access confidentiality because there is a chance of exploiting the information based on access frequency is increased. And thus it gets easy to find the content associated with the particular node. The shuffle index method discussed above is the best suit for giving the protection against above attack.

Content Confidentiality

The content security of the data is similar to the traditional security of the data. The data security is involved in the each and every stages of the lifecycle of data. Normally life cycle of data in cloud has five phases as

- Data generation
- Data transfer
- Data use
- Data share
- Data storage
- Archival
- Destruction

Figure 2: Data Life cycle in cloud

Data Generation: It is a process of generating the data which need to store on cloud.

Data Transfer: It is a process of transferring the data so data integrity and confidentiality of data is need to maintain.

Data use: In this stage the data is in clear format and not secured by the technique of csp encryption

Data Share: It is the process of sharing the data by the data owner, in advance they can share the same data to another parties where the integrity of data is maintained.

Data storage: This phase simply represents the way by which data is needed to store. It can be SaaS or PaaS.

Archival: It is the phase where the risk of data leakage is maintained.

Destruction: when the data is no longer required for futures then it is necessary to destruct the data as it is unnecessarily acquiring the space which in turn can be given to the another parties if required.

With this advantage the confidentiality of the data is at risk. Hence data confidentiality of cloud data is a hot topic under research. Numbers of techniques are being proposed to resolve the issues. The popular method of doing same is to encrypt the data. Since encryption gives high security and confidentiality, it is widely accepted by the indented entities. Still encryption is failed to provide the complete solution for the problem of access confidentiality because of which access confidentiality of the data becomes suspicious. In order to apply proper and effective encryption the algorithms used for the encryption and strength of the key should be considered. Since the Hugh amount of data is involved in cloud computing, the speed and transmission cost of Hugh encrypted data is need to consider. For doing so symmetric encryption is best suit compare to asymmetric encryption.

One of major problem of encryption is management of keys required for encryption. In general data owners have the rights of doing so, but due to the lack of expertise in the said domain they transferred there task to the cloud service providers. Since CSP maintains the Hugh number of keys it is merely difficult to maintain it.

Pattern Confidentiality

Pattern confidentiality is a process of giving protection to data against the capability of the server to recognize the two separate accesses to the individual node.

Data Distribution

One of the prominent features giving by cloud is data storage. Cloud owner never have the data on their own servers. Their data is located remotely and it is managed by cloud service providers. The main idea behind the distribution of data across the number of cloud service providers is to restrict the intruders for retrieving the data, as intruders can retrieve the data in and only if they have access to all of the nodes where data is being distributed. Even though the cloud service providers are claiming the privacy of the stored data, many breaches are found. So for this reason data owner cannot rely on single service providers. Hence reliability and integrity of cloud data can be easily achieved by fragmenting the data and storing across different physical locations. The described fragmentation can be done effectively by using four methods, as

1. Horizontal fragmentation
2. Vertical fragmentation
3. Mixed fragmentation
4. Derived fragmentation

Many techniques were proposed to fragment the data with less amount of encryption so that there will minimum amount of data exposure.

Numbers of algorithms are proposed for the purpose of encryption and decryption of data. E.g. AES, Reverse circle cipher, blowfish etc. Each of these encryption algorithms has their own advantage and disadvantage over other. Among all these algorithms reverse circle cipher emerged as a best suit for encryption and decryption as it requires less overhead to implement. Figure illustrates the working of this algorithm.

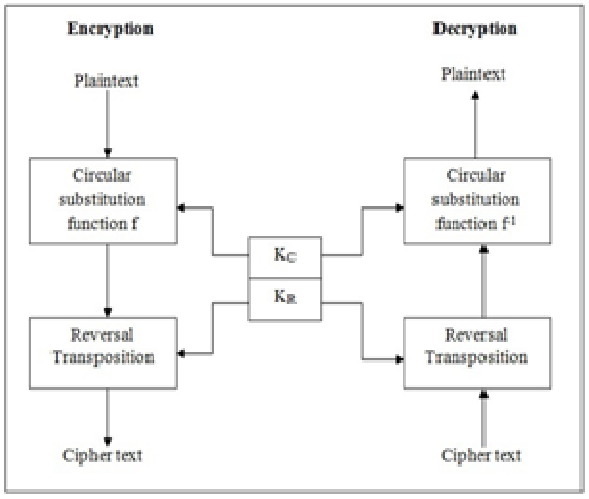

Figure 3: Reverse circle cipher algorithm

The rest of the paper is organized as follows. Section 2 discusses some related work and section 3 for conclusion.

II. LITERATURE SURVEY

This section represents all the related works of technologies used in our project.

As discussed above the data confidentiality and access confidentiality is at breach. [1] Provide the effective scheme that can be used for secure access to the outsourced data. To implement the idea concept shuffle indexes are used, while the B+ trees are used to incorporate the system. After testing the system with different scenarios authors find that the less cost is required to implement the system. Also they conclude that it is the first system to give the protection to the outsourced data while maintaining the content, access and pattern confidentiality.

[2] Proposed an efficient algorithm to maintain the security of the cloud data life cycle using the DIFFIE-HELMAN algorithm. One of the great things about the said system is that it can allow the two parties to communicate with each other and exchange there private keys over the channel which is completely unsecure for the transmission propose. Also system giving the protection against the intruders. Even if the intruders hack the data they are not able to decrypt the data and again the data is bring back to their original form. Since in this paper encryption/decryption schemes are not used which may become the bottleneck of the system so the authors keep this as a research work.

As discussed above shuffle index continuously rewrite and re-encrypt the data to deal with the issue of security of outsourced data. In shuffle index technique actual data is remain hidden from the external storage. In this method client can hide the actual request within the fake request and by doing so it shuffle the content within the block. The shuffling is done in such way that not only third parties but also the server is unable to find the link between the actual data and request data. Shuffle index algorithm is first proposed by the author Sabrina De Capitani di Vimercati and later on they started evaluating by other studies.

To overcome the problem of continuous re-writing and re-encryption [3] implements the concept of swapping where three independent servers are used to managed the data structures. Authors states that the main idea behind making use of 3 servers is security. The protection is giving in such way that for every request to access the node should be transfer to the different server. Also fail of any server will not harm the data. That's why such system always plays an important role in access confidentiality

To address the issue of cloud data distribution, [4] suggests an optimized method to distribute the data across different service providers available in the associated region. Here different schemes used for the fragmentations are highlighted over here. Also the comparison among the different fragmentation scheme is explained so it can be easy for the intended person to select the best one. When data is being distributed across the multiple csp security flaws are there. So to deal with the same

scenario [5] proposed a novel approach for the data fragmentation. The fragmentation scheme explained by author is for relational databases like MySQL where the numbers of tables are considers as fragments. While implementing the technique they taking care of amount of encryption being used. They are making minimum amount of encryption so that the less amount of data will get seen by the other parties.

Now a day's technique of swapping cloud data grabbing lots of attention as it keeps the data confidentiality. Swapping is a process of changing the server of data once it accessed by the users. Swapping creates confusion to the intruders as data keep on swapping continuously. In data swapping values of data are adjusted in such way that it swaps the fraction of records between the records so the third part intruders will not get actual data. Instead they got garbage data. Swapping methods are normally categorized in two sub parts.

1. Targeted swapping strategy
2. Random swapping strategy

[6] Gives data swapping technique to create confusion among the intruders. [7] Narrates a shuffling algorithm having distributed nature to maintain the access confidentiality of the cloud data. To accomplish the task, a new technic known as SHADOW is proposed by the authors. While proposing the method author have confidence that the system gives the more security with reduced cost. Here two servers are maintained to store the data and the view of each server is maintained in such way that there is only the single server have all the information. One of the major advantages of the system is that while preserving the security it is not degrading the performance of the system.

Dynamic data allocation locks the concurrency of the access blocks. Hence to overcome this problem shuffle index based techniques is used by [8] where all the said limitations are efficiently removed. In this paper concept of shuffle index is extended to support the various indexes and efficient concurrent access. Here both data and access privacy is maintained. [9] Explores a distributed architecture for storing the outsourced data to two servers which are completely untrusted, still maintain the data privacy. Also the partitioning is designed in such way that the content on any server will not have security bottlenecks. Also the step by step presentation is given which shows how to execute query on distributed architecture. And the challenges associated with systems are described.

[10] Implements the technique of multi rotational scheme used for maintaining strong security of the network data. Before this implementation, prior algorithms have the issue of complicated process of encryption, extra overhead of maintaining cost etc. Thus multi rotational scheme is proposed by the author who effectively covers all these issues. To make the system effective liner rotational scheme is boycott by them. In advance for the better understanding of the person who read it, they came with the detail implementation scheme. So one can follow it for implementation.

Coarse gain level encryption is one of the biggest problems observed in encryption algorithms. In coarse gain algorithms private key by using which encryption is done is forwarded to the intended user. So this technique has bottleneck of security. To overcome the same a new encryption scheme based on key policies is proposed by the writers [11] here labeling is done between the attribute and the key of the cipher text. Because of this system have full control on the user to which is going to decrypt the data. Also the author concludes that the proposed system has upper edge over the secret sharing schemes as in secret sharing scheme there should be coordination between the different parties where as this system does not require any kind of coordination. These concerns strongly increase the security of the system.

Sahai Explains the technique of single authority attribute based encryption; Due to few problems multi authority key based technique is remain as a future work of their solution. [12] took the Sahai's work as a base for their study and they overcome the problem of sahai by proposing multi authority based encryption scheme. In this scheme provision of monitoring the attribute and thus sharing of public keys is given to the any numbers of users. Prior data transferring data owner need to set the N value

which is the number of attributes. Once getting the ciphered data, receiver can decrypt it if and only if they have N same attributes. Hence authorized user should have the N same attributes. If not then they are not valid attributes.

Data partitioning is used to break down data into smaller chunks to manage and store it quickly without having overhead of storage management. Normally partitioning is performed in alphabetical manner by using certain set of indexes. It fetches the first two letters and checks whether the same folder with the fetched username is existed or not, if the folder is not existed then creates the folder and store file in that folder. This store file is in encrypted format with the key of encryption. When user wants to access the base file that time reconstruction of the partitioned file takes place to serve users. [13] Narrates the data partitioning scheme to increase the efficiency of the cloud data storage. System effectively reduces the cost of storage and thus reducing the time complexity of the system. Apart from this dynamic operation scheme is proposed where secure encryption and decryption operations are takes place. One more plus point of the system is that it checks the integrity of the stored data to deal with the third party intruders.

[14] Elaborates the very first approach to perform partitioning of data in mobile cloud computing system. As per the research concern this is the first algorithm to focus on the partitioning of cloud data. In first phase of the system dynamic partitioning and execution of application is proposed. Unlike the prior systems here apart from dynamic partitioning, system have provision of sharing instances to the multiple users which in turn speed up the process. As the design of the framework is based on the elastic cloud it gives higher scalability. To show the efficiency of the system both real word and numeric computations are performed, which shows that the proposed system is giving the better result over the other.

Figure 4 illustrates the working of data stream in mobile applications.

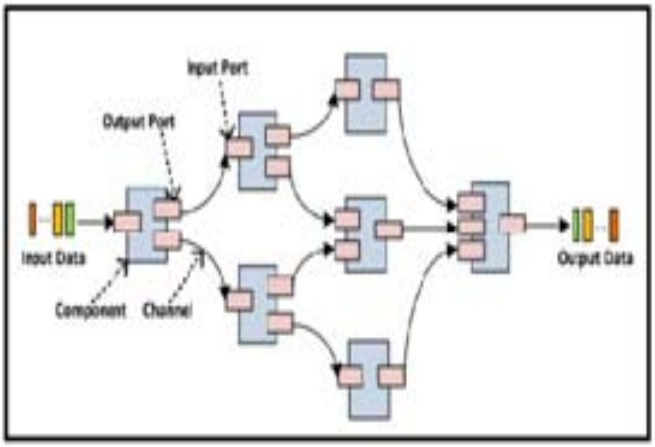

Figure 4: The Model for Data Stream Applications

[15] Presents a locality aware and fairness aware scheme for key partitioning in map reduce which named as LEEN. Here all the intermediate keys those are buffered are partitioned according to their number of occurrences. To perform experimental evaluation system is implemented on Hadoop -0.18.0 which shows that the system gives the higher locality and thus it reduces the amount of data being shuffled. Authors observe Up to 40% improvement when system is tested on the different workloads. In future they wanted to integrate the LEEN plugin in Hadoop and want to evaluate the system for the different domains such as scientific applications. To improve the query based load balancing there is need to develop the same system for the query optimization.

In order to take complete benefit of cloud, the way by which data stored in cloud should be managed properly. To obtain this an effective indexing scheme is need to implement. This indexing scheme is responsible for maintaining low cost and speed up the process of searching by incorporating the parallel search. [16] Implements a B tree based scheme for indexing of cloud data. Here CG index (cloud global index) scheme is proposed which is a secondary indexing scheme of cloud data storage. The maintenance of indexing technique is done in incremental way.

Figure 5 explains the B tree nodes and the Associated index ranges.

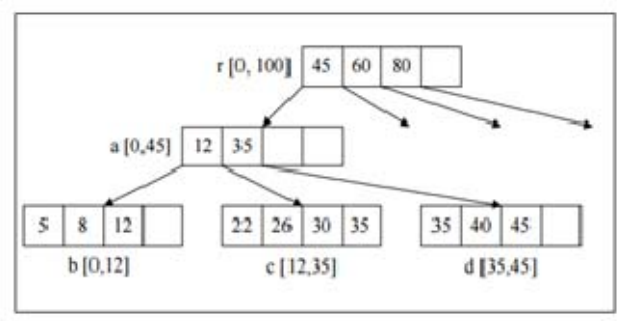

Figure 5: Representation of B tree nodes and indices

III. Conclusion

The protection of the confidentiality of outsourced data is an important problem. A critical aspect is the ability to efficiently access data that are stored in an encrypted format, without giving to the server managing access requests the ability to infer knowledge about the data content of the access executed by the clients. The approaches that have been proposed to solve this problem rely on a continuous rewriting and re-encryption of the accessed data, like the shuffle index that has recently been proposed.

So many methodologies are discussed in this paper which gives overall idea about data swapping to provide more confidentiality in cloud.

Acknowledgment

Author is sincerely grateful to Prof. Sonali Patil, my Project guide and mentor for her valuable guidance and encouragement. Also the authors are thankful to the Computer Engineering Department of JSPM's Bhivarabai Sawant Institute of Technology & Research for their support in providing a good environment and facilities like books, internet and the other resources to complete this research.

References

1. De Capitani di Vimercati, Sabrina, et al. "Efficient and private access to outsourced data." Distributed Computing Systems (ICDCS), 2011 31st International Conference on. IEEE, 2011
2. Rautela, Sangita, Arvind Negi, and Prashant Chaudhary. "Data Security and Updation of Data Lifecycle in Cloud Computing using Key-Exchange Algorithm."
3. De Capitani di Vimercati, Sabrina, et al. "Protecting access confidentiality with data distribution and swapping." Big Data and Cloud Computing (BdCloud), 2014 IEEE Fourth International Conference on. IEEE, 2014.
4. Reddy, B. AmarNadh, and P. Raja Sekhar Reddy. "Effective Data Distribution Techniques for Multi-Cloud Storage in Cloud Computing." CSE, Anurag Group of Institutions, Hyderabad, AP, India.
5. Hudic, Aleksandar, et al. "Data confidentiality using fragmentation in cloud computing." Int. J. Communication Networks and Distributed Systems 1.3/4 (2012): 325-329.
6. Barik, Sachida Nanda. "Data Swapping in Cloud Computing." (2015).
7. di Vimercati, Sabrina De Capitani, et al. "Distributed shuffling for preserving access confidentiality." Computer Security–ESORICS 2013. Springer Berlin Heidelberg, 2013. 628-645.
8. di Vimercati, Sabrina De Capitani, et al. "Supporting concurrency and multiple indexes in private access to outsourced data." Journal of Computer Security21.3 (2013): 425-461.
9. Aggarwal, Gagan, et al. "Two can keep a secret: A distributed architecture for secure database services." CIDR 2005 (2005).
10. "Enforcing Reverse Circle Cipher for Network Security Using Multirotational Technique", Sajjade Zeba S. International Journal of Advanced Research in Computer Science and Software Engineering
11. "Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data" Vipul Goyal¤ Omkant Pandey Amit Sahai Brent Waters
12. "Multi-Authority Attribute Based Encryption" Melissa Chase Computer Science Department Brown University Providence, RI 02912
13. Swapnil V.Khedkar et al, "Data Partitioning Technique to Improve Cloud Data Storage Security."(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3347-3350
14. Yang, Lei, et al. "A framework for partitioning and execution of data stream applications in mobile cloud computing." ACM Sigmetrics Performance Evaluation Review 40.4 (2013): 23-32.

15. Ibrahim, Shadi, et al. "Leen: Locality/fairness-aware key partitioning for mapreduce in the cloud." Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on. IEEE, 2010.

16. Wu, Sai, et al. "Efficient B-tree based indexing for cloud data processing. “Proceedings of the VLDB Endowment 3.1-2 (2010): 1207-1218.

AUTHOR(S) PROFILE

Ashwini Taksal, received the B.E. degree in Computer Science and Engineering from Bhivarabai Sawant Institute Of Technology and Research in 2014. Currently she is pursuing her M.E. in Computer Science and Engineering under the guidance of Prof. Sonali A. Patil from Bhivarabai Sawant Institute of Technology and Research from Pune, India.

Prof. Sonali A. Patil, received the B.E. and M .Tech. degree in Computer Science and Engineering and Currently pursuing PHD. She is currently working as an Assistant Professor in JSPM’s Bhivarabai Sawant Institute Of Technology and Research, Wagholi, Pune, India.

