Indian Journal of Science and Technology, Vol 9(21), DOI: 10.17485/ijst/2016/v9i21/95239, June 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Cloud Computing Security with Collaborating Encryption

N. D. Jambhekar[1*], Sanjay Misra[2] and C. A. Dhawale[2]

[*1]Department of Computer Science, S.S.S.K.R. Innani Mahavidyalaya, Karanja(Lad), Washim - 444105, Maharashtra, India; ndjambhekar@rediffmail.com
[2]Covenant University, Nigeria; ssopam@gmai.com
[3]P.R. Pote College of Engineering & Management, Amravati - 444602, Maharashtra, India; cadhawale@rediffmail.com

Abstract

The security of bigger data is the bottleneck in the encryption and decryption because of the big data size. The single encryption technique using one source is not adequate to accomplish the big data cloud computing security. This paper elaborates the working of cloud computing and the collaborating source security system for the Big Data security. A collaborating encryption technique framework is proposed in this paper to meet the futures' faster encryption requirements. The traditional information security system is not capable to provide the complete security during the cloud computing. The method described in this research comprises the parallel and distributed encryption system which gets the benefits the homomorphic encryption technique. The encryption facility during the mobile communication of object is tedious. Every cloud has its own security features and can be working in collaboration with the other cloud servers. Therefore, the parallel and distributed encryption facilities can be possible at every next door of other cloud without breaking the sequence of encryption process. The essential resources become the remote resources and the allocation of these resources can be managed at every cloud. Most of the time while working with the cloud computing is the availability of network and other resources. Providing the information security in the unavailability of resources during encryption and decryption is a difficult task. The collaboration encryption technique is a framework where, different clouds can work in parallel with the distributed processing. The security mechanism is improved by the homomorphic encryption.

Keywords: Big Data, Cloud Computing, Collaborating Encryption, Encrypted Search

1. Introduction

1.1 Cloud Computing

Cloud computing plays a vital role in mobile computing by offering on demand network and resource access within a shared pool independent ready to use mobile resources such as networks, collaborating servers, virtual data storages, tools and applications without any effort[1]. The cloud is a mobile framework that contributes the availability different services and resources with different cloud models. Figure 1 depicts the cloud computing and data storage scenario. The services provided by different clouds are-

1.1.1 On-demand Self Service

Automatic services available from the cloud to anyone connected to the cloud.

1.1.2 Heterogeneous Platforms

Network access can possible to any type of devices having any platform without cloud-to-cloud surfing[2].

1.1.3 Resource Sharing

A pool of resources is available on every cloud and scheduled for the connected users while moving from one

*Author for correspondence

location to another[3]. Cloud physical resources acts like virtual and are location independent[4]. When the user moves from one cloud to another, these resources are available on every cloud without specifying their locations. These resources include virtual storage, processors, network bandwidth[5].

1.1.4 Cloud Collaboration

Multiple clouds are working in collaboration to provide the uninterrupted services to the connected users. The movements from one location to another do not affect the mobile computing of the device[6–8]. The cloud services work in collaboration and available to the user moving from one cloud to another.

Different types of clouds are available to the users such as

- Private Cloud: It is an enterprise cloud available for the campus area defined users with internal access to resources and services solely on demand.
- Community cloud- It is a backbone network of clouds shared by multiple organizations, having their limited own policies.
- Public Cloud- The public cloud is accessible to anyone with all the resources and functionalities available to the connected user publicly. Sometimes it is owned by a public organization and the users requested the cloud services on demand.
- Hybrid cloud- More than one type of cloud works in collaboration to provide the services to the users moving from one location to another and without demanding the services from cloud to cloud.

1.2 The Big Data

Big data is an emerging technology that comparatively gain its important and even more than the cloud computing. Cloud computing revolutionaries the world of virtual smart computing and data storage.

Big data is a massive set of information comprised of different data type having dissimilar data structure likely to be stored in multiple locations[9]. Every day, huge amount of information is captured and disseminated to and from the remote locations and requires massive amounts of storage area. Due to the extraordinary monolithic structure, the data can be difficult to organize, store, analyze, and retrieve[10–12].

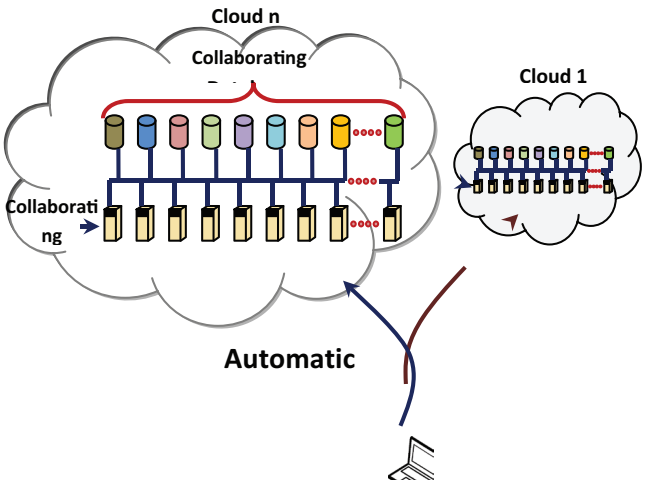

Figure 1. Cloud computing & data storage scenario.

Today, the world moving towards the digital age. Every hard coded paper document will be transformed into digital form. The digital libraries take the place of traditional paper book libraries. Keeping the massive data in a paperless digital form is simpler and easier than keeping it in the form of hard coded paper such as keeping in the paper libraries[13]. However, this simplicity to store the huge amount of information in digital form requires a large number of digital manipulating resources[14]. Though it solves the problem of storage space required for the paper libraries, it can provide other benefits also such as easy to access and manipulate the digital information.

1.2.1. Big Data Security Challenge

Big data is a massive set of digital information stored on different servers and can be available through the cloud[15]. If data are smaller, it can easily encrypt/decrypt during transmission and storage. The bigger data suffer from the security problems, hardware offload, operating system and resources management, data acquisition, data analysis and processing workloads, indexing, cataloging, searching, data mining and dissemination[16]. If the data is in text form it is sometimes easier, but if it is audio, images and video data, then it is harder to store and manipulate because of the large size. If data captured by any organization in a few terabytes or petabytes size, it is harder to process by a single machine even by the supercomputer or cluster machines[17–19]. Following Figure 2 depicts the security scenario along with the big data storage and the required resources within the cloud.

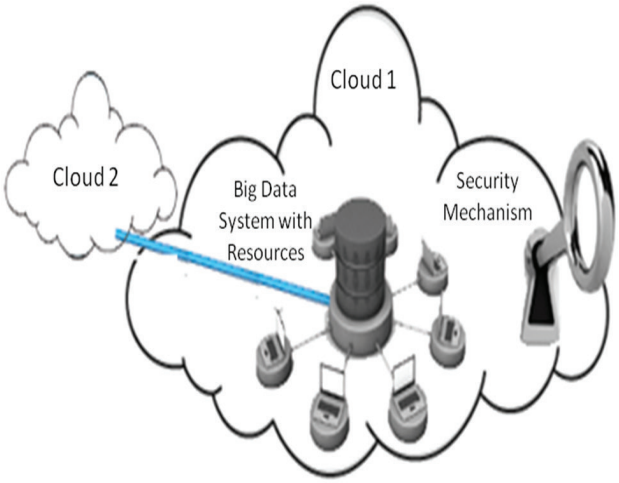

Figure 2. Big Data security.

1.2.2 Encrypted Storage

Every organization or an individual trying to save their confidential information over cloud storage. A massive amount of data is captured by the cloud from various resources where it may be a secured encrypted data or plain confidential information[20]. The client computer can implement the security or the cloud server or both can be engaged to encrypt the data. As the data larger, it is harder to encrypt by a single cloud server and store it[21].

1.2.3 Encrypted Workload

Workload is the amount of data handle during the encryption process. If the incoming data becomes larger, the encryption workload becomes heavier. The client or server can handle the encryption process. If the client performs the encryption and server becomes free, but the network traffic suffering due to transfer of large encrypted data[22]. For several security reasons, it is better to encrypt the data at its source i.e. at the client machine[23]. It can resolve the problem of encryption, decryption, key maintenance and transfer.

1.2.4 Decryption

If the cloud server store the data encrypted by clients, then the client machine can handle the part of decryption. Only the work is to transfer the encrypted data towards the client computer.

1.2.5 Encrypted Storage

As the data are bigger, the retrieval is tedious because the indexing, ordering and searching consumes lot of computing and network resources.

1.2.6 Failure and Recovery

The Database Management System (DBMS) uses the ACID properties, for example, Atomicity, Consistency, Isolation, and Durability to work with the transactions. It assures the complete successful transaction processing. If the failure occurs, the log based recovery system is available.

In big data cloud computing, the recovery can be implemented by the same log based concept with the implementation of ACID properties.

This paper investigates the current security issues of big data and cloud computing and proposed a framework to enhance the security. The essence of big data security is not handled by a single framework, rather it can be efficiently controlled by the collaborating encryption system where the entities engage in this big data security can be collaborating cloud servers working with the client security mechanism.

2. Existing Big Data Security Framework

Recently, the following big data security challenges strike

2.1 Secure Parallel and Distributed Processing

The big data from the client is divided into equal number chunks, and processed in the parallel and distributed way. As the data is divided for encryption and collected again, the security at different level from machine to machine must be preserved[24–26].

2.2 Secured Data Storage and Retrieval

However, the size of the data storage increases exponentially from cloud to cloud, the scalability and availability is a major challenge while maintaining the huge data.

2.3 Source Input Validation

Voluminous data comes to store from a variety of sources. How can it be possible to ensure the data incoming from the trusted source? Here the input validation is necessary and is a big challenge to control the data storage from the valid data sources.

2.4 Active Monitoring

The big data real time active security monitoring is a major challenge. As the data is bigger, several cloud

servers are required to monitor the real time data flow for other cloud servers are storages.

2.5 Privacy Preserving

As the data is bigger, it must store on different cloud servers. The transaction log can also store on the same servers that plays a big role in privacy leakage. It is a major issue to restrict the private storage and retrieval of the confidential information over big data cloud storage system. The real time activities leak the privacy.

2.6 Secure Communication

The data storage on the cloud by the client has to be cryptographically secured before transferred over the insecure communication channel.

2.7 Access Control

No access other valid sources must be restricted to preserve the security of the data. The access to the original source by identifying its authentication restricts the malicious users. This can be possible by maintaining the metadata about the user and their access.

3. Proposed Solution to Big Data Security

As the data is bigger, the security is harder. Weaker security is a bottleneck for the cloud computing and big data technology. The following section explores the big data platform.

3.1 Big Data Platform

Not all the data is useful and scientist working to refine the needful data across the big data. An infrastructure is required for the Big Data that can store, move, and integrate huge data with greater speed and accuracy where the traditional infrastructure not satisfy this need. The way is to transform the unstructured Big Data into a structured form with the help of complex structured database management system. Figure 3 depicts the required Big Data security platform.

- Storage: Large storages are required to store the huge voluminous data. As the user increases, the larger storage required. Not a single store point keeps all the data; therefore, the parallel collaborating storages are required for the big storage.
- Communication and Distribution: Voluminous data is moved from client to Big Data cloud storage and server to server requires great communication and distribution capacity with speed and agility.
- Structuring of unstructured data: The incoming data are unstructured and require to transform into a structured form requires larger hardware, processing and networking cost. The structure data is useful and unstructured data occupy unnecessary storage space.
- Metadata management: As the data is bigger, the metadata also becomes bigger. The unstructured data cannot yield the right metadata. To search the useful data across the metadata, the data required be structured.

3.2 Big Data Processing

- The cost to process the huge data is heavier. The hardware, storage, softwares and networking cost is larger to process the bigger data.
- Sharing: Data sharing across client to servers, servers to servers and cloud to cloud is a time consuming task. The sharing of data introduces a security trouble. The secrecy of the confidential data is a prime goal for every organization, individual and the Big Data cloud itself.
- Transition: The unstructured data conversion into structured for solving the trouble of indexing, ordering and recognizing the useful content of the data.
- Retrieval: The complex query structure is required to retrieve the useful contents across the huge data storage[27]. The well-organized and well-structured data is easier for searching and save the cloud resources. As the users increases, the retrieval, communication and sharing speed declines.
- Query processing and Views: Complex query processing efficiently deals with the huge data that produces the fruitful views of the clients. The distributed query processing with collaborating functions from multiple servers across

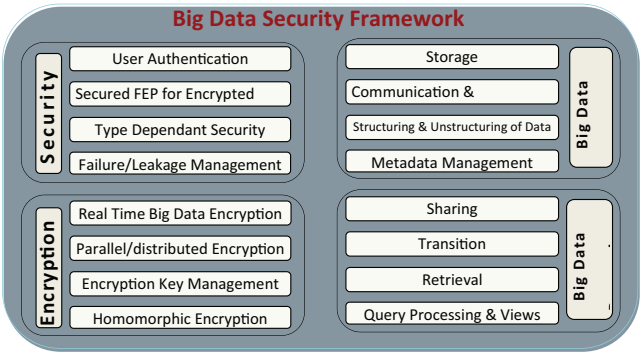

Figure 3. Big Data security platform.

multiple clouds and their storages helps to yield the valuable data across the huge data.

3.3 Security

- User/Administrator Authentication: The registered users, guest, administrator authentication requires via preserving the metadata for them. The security is controlled by a software or a hardware base Front End Processor for every cloud big data.
- Secured Front End Processors for encrypted search: Encrypted search is a major essential part of a Big Data security system. The software or hardware base front-end processors play a great role in maintaining the security across the cloud and Big Data.
- Type dependant security: Different type of data requires different security mechanism. The well-known cryptographic algorithms such as AES carry out the text data security. While the audio, images and video security is carried out by the various steganographic techniques.
- Failure/Leakage Management: The log-based recovery must be maintained for the failure. The Two-Phase-Locking protocol is useful in Big Data leakage management.

3.4 Encryption

Encryption plays a bigger role in the security of Big Data information. Various encryption techniques are available such as AES, TDES to secure the data.

- Real time big data encryption: The offline encryption is suitable for the small local data. However, the huge online data must be secured by the real time encryption. This can be possible only by capturing the data its source and encrypt. The client encrypts the data and transferred over cloud solves the key management problems.
- Parallel/Distributed encryption: If cloud carries the security work huge data, it is well handled by the parallel and distributed encryption technique. Multiple servers across the different clouds perform the encryption in collaboration.
- Encryption Key Management: A trusted third party plays a great role to maintain the encryption key. If user encrypts the data at its source, the key need not be managed at server. If servers handle the encryption part, the digital certificate technique is useful.
- Homomorphic Encryption: A newer fully homomorphic encryption technique used to merge different parts of encrypted data, which supports the encrypted search. The homomorphic encryption plays a role in multiparty encryption in keeping the encrypted data at combining place[28,29]. The collaborating encryption is helpful for the homomorphic encryption where clients from different places encrypt their own data and store at a single cloud. The query in encrypted form can be searched on that server without understanding the meaning of the ciphertext.

4. Conclusion

The practical implementation of security framework for the cloud computing big data is somewhat tedious. The present cryptosystem is not enough useful for full security. If the traditional encryption techniques are applied, then the encrypted data stored across the cloud big data server storages cannot search at all. The fully homomorphic encryption technique is a miraculous solution and can support the encrypted search. This can be implemented with the help of the collaborating encryption technique discussed in this paper. This paper also focuses on the real time encryption for the flowing data over the cloud. This is possible with the help of collaborating servers among cloud to cloud. Major technology variations occur in a few years where traditional encryption techniques not fulfill the security need. In this paper, the parallel and distributed encryption technique introduced that helps to reduce the confidential data security require-

ment. It helps to encrypt the parts of data at different locations and merge at a single location.

5. References

1. Ranjan R. Streaming big data processing in datacenter clouds. Proceedings of IEEE Cloud Computing. 2014 May; 1(1):78–6.
2. Bagheri R, Jahanshahi M. Scheduling workflow applications on the heterogeneous cloud resources. Indian Journal of Science and Technology. 2015 Jun; 8(12):1–8. DOI: 10.17485/ijst/2015/v8i12/57984.
3. Zhao F, Li C, Liu CF. A cloud computing security solution based on fully homomorphic encryption. Proceedings of 16th International Conference on Advanced Communication Technology (ICACT); 2014 Feb. p. 485–4.
4. Jasmine RM, Nishibha GM. Public cloud secure group sharing and accessing in cloud computing. Indian Journal of Science and Technology. 2015 Jul; 8(15):1–7. DOI: 10.17485/ijst/2015/v8i15/75177.
5. Jeuk S, Szefer J, Zhou S. Towards cloud, service and tenant classification for cloud computing. Proceedings of 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid); 2014 May. p. 792–10.
6. Murthy PK. Top ten challenges in Big Data security and privacy. Proceedings of IEEE International Test Conference (ITC). 2014 Oct. p.1.
7. Kirubakaramoorthi R, Arivazhagan D, Helen D. Analysis of cloud computing technology. Indian Journal of Science and Technology. 2015 Sep; 8(21):1–3. DOI: 10.17485/ijst/2015/v8i21/79144.
8. Pal AS, Pattnaik BP. Classification of virtualization environment for cloud computing. Indian Journal of Science and Technology. 2013 Jan; 6(1):127–33. DOI: 10.17485/ijst/2013/v6i1/30572.
9. Xu L, Jiang C, Wang J, Yuan J, Ren Y. Information security in big data: privacy and data mining. IEEE Access. 2014 Oct; 2:1149–28.
10. Rajathi A, Saravanan N. A survey on secure storage in cloud computing. Indian Journal of Science and Technology. 2013 Apr; 6(4):4398–401. DOI: 10.17485/ijst/2013/v6i4/31871.
11. Lee JY. A study on the use of secure data in cloud storage for collaboration. Indian Journal of Science and Technology. 2015 Mar; 8(S5):33–6. DOI: 10.17485/ijst/2015/v8iS5/61462.
12. Parthiban P, Selvakumar S. Big data architecture for capturing, storing, analyzing and visualizing of web server logs. Indian Journal of Science and Technology. 2016 Jan; 9(4):1–9. DOI: 10.17485/ijst/2016/v9i4/84173.
13. Ahmed ST, Loguinov D. On the performance of mapreduce: a stochastic approach. Proceedings of IEEE International Conference on Big Data (Big Data); 2014 Oct. p. 49–54.
14. Hu H, Wen Y, Chua TS, Li X. Toward scalable systems for big data analytics: a technology tutorial. IEEE Access. 2014 Jun; 2:652–36.
15. Chen XW, Lin X. Big data deep learning: challenges and perspectives. IEEE Access. 2014 May; 2:214–12.
16. Matturdi B, Xianwei Z, Shuai L, Fuhong L. Big data security and privacy: a review. China Communications. 2014 Supplement; 11(14):135–11.
17. Ren DQ, Wei Z. A failure recovery solution for transplanting high-performance data-intensive algorithms from the cluster to the cloud. Proceedings of IEEE International Conference on High Performance Computing and Communications & IEEE 10th International Conference on Embedded and Ubiquitous Computing (HPCC & EUC); 2013 Nov. p. 1463–6.
18. Singh J. Real time BIG data analytic: security concern and challenges with machine learning algorithm. Proceedings of Conference on IT in Business, Industry and Government (CSIBIG); 2014 Mar. p. 1–4.
19. Alguliyev R, Imamverdiyev Y. Big data: big promises for information security. Proceedings of IEEE 8th International Conference on Application of Information and Communication Technologies (AICT); 2014 Oct. p.1–4.
20. Hwang YH, Seo JW, Kim IJ. Encrypted keyword search mechanism based on bitmap index for personal storage services. Proceedings of IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom); 2014 Sep. p. 140–8.
21. Kalpana V, Meena V. Study on data storage correctness methods in mobile cloud computing. Indian Journal of Science and Technology. 2015 Mar; 8(6):495–500. DOI: 10.17485/ijst/2015/v8i6/70094.
22. Marchal S, Jiang X, State R, Engel T. A big data architecture for large scale security monitoring. Proceedings of IEEE International Congress on Big Data (BigData Congress); 2014 Jun 27–Jul 2. p. 56–8.
23. Dong X, Li R, He H, Zhou W, Xue Z, Wu H. Secure sensitive data sharing on a big data platform. Tsinghua Science and Technology. 2015 Feb; 20(1):72–9.
24. Tan Z, Nagar UT, He X, Nanda P, Liu RP, Wang S, Hu J. Enhancing big data security with collaborative intrusion detection. IEEE Cloud Computing. 2014 Sep; 1(3):27–7.
25. Bosch C, Peter A, Leenders B, Lim HW, Tang Q, Wang H, Hartel P, Jonker W. Distributed searchable symmetric encryption. Proceedings of Twelfth Annual International Conference on Privacy, Security and Trust (PST); 2014 Jul. p. 330–8.
26. Shyamala K, Rani TS. An analysis on efficient resource allocation mechanisms in cloud computing. Indian Journal of Science and Technology. 2015 May; 8(9):814–21. DOI: 10.17485/ijst/2015/v8i9/50180.

27. Ji C, Li Y, Qiu W, Awada U, Li K. Big data processing in cloud computing environments. Proceedings of 12th International Symposium on Pervasive Systems. Algorithms and Networks (ISPAN); 2012 Dec. p. 17–7.
28. Chan ACF. Symmetric-key homomorphic encryption for encrypted data processing. Proceedings of IEEE International Conference on Communications ICC '09; 2009 Jun. p. 1–5.
29. Xiang G, Yu B, Zhu P. A algorithm of fully homomorphic encryption. Proceedings of 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD); 2012 May. p. 2030–4.