International Journal of Trend in Scientific Research and Development (IJTSRD)
International Open Access Journal
ISSN No: 2456 - 6470 | www.ijtsrd.com | Volume - 2 | Issue – 5

Secured File Storage System In Big Data With Cloud Access Using Security Algorithms

U. Prathibha, Dr. G. Anitha, J. Ramyabharathi
Assistant Professor, Department of Computer Applications,
Karpagam Academy of Higher Education, Tamil Nadu, India

ABSTRACT
Big data is a technology to huge data sets, have high Velocity, high Volume and high Variety and complex structure with the difficulties of management, analyzing, storing and processing. The paper focuses on extraction of data efficiently in big data and how to manage the data and the components that are useful in handling big data. Security in the era of big data and especially to the problem of reconciling security and privacy models by exploiting the map reduce framework. Data can be classified as public, confidential and sensitive This paper proposes the big data applications with the Hadoop Distributed Framework for storing huge data in cloud in a highly efficient manner In order to avoid the third party issues and produce the exact data to the user by implementing the encryption and decryption approach using SHA 512 algorithms to avoid the security issues in big data.

Keywords: Big data tools, Hadoop, HDFS, Map Reduce, Encryption and Decryption Algorithms, SHA 512 Algorithms

I. INTRODUCTION
Big data means actually a big data; it is a collection of huge datasets that cannot be handled using old computing techniques. Big data is not only containing data, it also contains various tools, techniques and frameworks. Data that has extra-large Volume, comes from Variety of sources, Variety of formats and comes at us with a great Velocity is normally referred to as Big Data. Big data can be structured, unstructured or semi-structured. Big data hold the data generate by various equipment and applications like Black box Data which is a part of helicopter. In the proposed system the encryption files to be uploaded in the cloud. The integrity and confidentiality of the data uploaded by the user is ensured doubly by not only encrypting it but also providing access to the data only on successful authentication. The existed file on the device will be encrypted using SHA 512 algorithm. To enhance security; key will be encrypted using SHA 512 algorithm and will be stored in intern server. The authorized user can also download any of the uploaded encrypted files and read it on the system.

II. LITERATURE REVIEW
In [1] Abid Mehmood and Iynkaran Natgunanathan 2016 explained a comprehensive overview of the privacy preservation mechanisms in big data and presented the challenges for existing mechanisms. It illustrated the infrastructure of big data and the state-of-the-art privacy-preserving mechanisms in each stage of the big data life cycle. In [2] Ms. Chetana Girish Gorakh and Dr. Kishor M. Dhole 2016 presented an overview of big data concepts and characteristics and discussed introductory scenario about tools used in big data environment. It also covers security issues for big data. It discussed the huge data configuration, distribution and analysis that overcome the drawbacks of traditional data processing technology to manage, store and acquire data very speedily and cost effective, involves various tools, technique and framework. In [3] G Geethakumari and Agrima Srivatsava, 2012 explained to investigate the impact of Big Data techniques when applied in the field of Enterprise data Security and to develop the analysis and design techniques to mitigate the security threats so as to secure the Enterprise data more efficiently.In [4] Kalyani Shirudkar and Dilip Motwani 2015 described different security methods like Type Based keyword search for security of big

data, use of hybrid cloud to provide privacy in big data and also represented the application of big data in malicious url filtering.. In [5] Khalid Adam Ismail Hammad and Pahang Kuantan et.al., 2015 discussed about a more attention to cloud computing and Big data focused on "data", like data service, data acquisition, analysis and data mining, which pays more attention on ability of data storage and also discussed big data modeling and big data security issues. In [6] Nirali Honest and Atul Patel, 2016 described the limitation of traditional approach to manage the data and the components that are useful in handling big data and used Hadoop framework with the major components of the framework and working process within the framework. In [7] O. Liu and K.L. Man et.al, 2016 discussed the mechanism of preferential attachment during network evolution, which is considered one of the key factors in the formation of scale-free networks and tested the effectiveness of this model by a simulation using data of a real-world Chinese social network. Here they used the advanced analytics, enterprises can analyzed big data to learn about relationships underlying social networks that characterize the social behavior of individuals and groups.In [8] Prachi Pardeshi and Komal Patil,et.al 2016 proposed MPBTM model to generate pattern enhanced topic representations to model user's interests across multiple topics. In the filtering stage, the MPBTM selected the maximum matched patterns, instead of using all discovered patterns, for estimating the relevance of incoming documents. In this approach incorporates the semantic structure from topic modeling and the specificity as well as the statistical significance from the most representative patterns. In [9] Priyank Jain and Manasi Gyanchandani , et.al , 2016 described the recent techniques of privacy preserving in big data like hiding a needle in a haystack, identity based anonymization, differential privacy, privacy-preserving big data publishing and fast anonymization of big data streams and explained the Comparative study between various recent techniques of big data privacy is also done as well. In [10] Rishabh Mishra and Dr. Rakesh Sharma 2015 explained about the significance and opportunities of big data to identified from different perspectives and described the grand challenges (namely, data complexity, computational complexity, and system complexity), as well as possible solutions to address the forthcoming challenges. In [11] Samiddha Mukherjee and Ravi Shaw 2016 discussed the future opportunities that could be harnessed in this field, where much of the research is yet to be done and described the hurdles of securing the data and democratizing it have been elaborated amongst several others such as inability in finding sound data professionals in required amounts and software that possess ability to process data at a high velocity. In [12], Shantanu Kalbhor and Hiteshkumar Jain et.al, 2016 explained that big data analytics is the procedure of study huge amounts of data and represented by the dimensions volume, variety, velocity and veracity. Hadoop framework with map reduce paradigm to process the data. In [13], Sharifnawaj Y. Inamdar et al, 2016 defined the Hadoop to provide the distributed storage and Elastic Map Reduce for the clients to run their jobs. They demonstrated that Hadoop is a key to execute the security of client information in such systems.In [14] SHASHANK , S.K.Saravanan, G.Rekha , 2015, described effective use of big data requires access from any domain to data in that domain, or any other domain it is authorized to access. In [15] Sophia Yakoubov and Vijay Gadepally et.al described the cryptographic techniques in the context of our cloud model and highlight the differences in performance cost associated with each. In [16] Srinivisan Nagaraj and Kishore Bha midipati et.al 2010 explained the file security three different kinds of algorithms i.e. RIJNDAEL, Initialization vector (IV), SHA512 Hashing algorithm these three popular algorithms are used for file encryption and decryption approach. In [17] Venkata Narasimha Inukollu and Sailaja Arsi et.al 2014 discussed the security issues in cloud computing that are associated with big data and used the various big data tools for cloud environments can be secured for complex business strategies. In [18], Zaid KARTIT and Mohamed EL 2015, proposed a simple, secure, and privacy-preserving architecture for interCloud data sharing and is based on an encryption/decryption algorithm which aims to protect the data stored in the cloud from the unauthorized access. In [19] Zhang Hongjun and Hao Wenning 2014 discussed the characteristics of big data information security, and focused on conclusion of security problems under the big data field and the inspirations to the development of information security technology.

III. BIG DATA

Big data technology is used to store huge amount of data. It is an innovative technology which is more applicable for both academics and corporate sectors. Security and privacy issues are determined by the

Volume, Variety, Velocity, Value and Veracity of big data

Volume - Volume defined to store huge volume of data can access to data and perform operations.

Variety - Variety that deals with the different types of data from various sources that big data frameworks

Velocity - Velocity refers that big data systems can retrieve and store data independently at different rates in which flow of data may occur in or out the system and provides an abstraction layer

Value- Value defines the exact value of data i.e., the efficient value of the data for given information. If they are not provided an exact value otherwise all the data are worthless.

Veracity -Veracity consider the trustworthy data, addressing data confidentiality, data integrity, and data availability.

I. Challenges with Big Data

A. Heterogeneity and Incompleteness

If we want to evaluate the data, it should be structured but when we good deal with the Big Data, data may be structured or unstructured as well. Heterogeneity is the big challenge in data Analysis and analysts need to cope with it. Consider an example of patient in Hospital. We will generate each record for each medical test. And we will also generate a record for hospital stay. This will be unequal for all patients. This design is not well structured. So managing with the Heterogeneous and incomplete is required. A good data analysis should be applied to this.

B. Privacy

Privacy of data is one bigger problem with big data. In some countries there are tough laws about the data privacy, for example in USA there are tough law for health records, but for others it is less forceful. For example in social media we cannot get the private posts of users for sentiment analysis

C. Scale As the name says Big Data is having huge size of data sets.

Managing with large data sets is a big problem from decades. In previous years, this problem was solved by the processors getting faster but now data quantity is becoming large and processors are static. World is moving towards the Cloud technology, due to this shift data is generated in a very high rate. This high rate of increasing data is becoming a challenging problem to the data analysts. Hard disks are used to store the Data. They are slower I/O performance. But now Hard Disks are replaced by the solid state drives and other technologies. These are not in slower rate like Hard disks, so new storage system should be constructing.

D. Human Collaborations

In spite of the advanced computational models, there are many patterns that a computer cannot recognize. A new method of harnessing human ingenuity to solve problem is crowd-sourcing. Wikipedia is the perfect example. We are reliable on the data given by the strangers, however most of the time they are correct. But there can be other people with other motives as well as like providing false data. We need technological model to handle with this. As humans, we can look the review of book and find that few are positive and few are negative and come up with a decision to whether buy or not.

II. Opportunities to Big Data

A. Media

Media is using big data for the boost and sale of products by focus the interest of the user on internet. For example social media posts, data analysts get the number of posts and then evaluate the interest of user. It can also be done by taking the positive or negative reviews on the social media.

B. Government

Big data can be used to handle the issues faced by the government. Obama government declared big data research and development initiative in 2012. Big data analysis played an important role of BJP winning the elections in 2014 and Indian government is implement big data analysis in Indian electorate. 4.3. Technology.

Almost each top organization like Facebook and yahoo has adopted Big Data and are spending on big data. Facebook holds 50 Billion photos of users. Every month Google holds 100 billion searches. From these stats we can say that there are a lot of opportunities on internet, social media.

C. Science and Research

Big data is an up-to-the-minute topic of research. Large number of Researchers is working on big data. There are so many papers being published on big data.

D. Healthcare

According to IBM Big data for Healthcare, 80% of medical data is unstructured. Healthcare organizations are adapting big data technology to grab the complete data about a patient. To boost the healthcare and low down the cost big data analysis are needed and certain technology should be adapted.

IV. BIGDATA TOOLS

Big Data is a term used for a collection of data sets so large and complex that it is difficult to process using traditional applications/tools. It is the data exceeding Terabytes in size. Because of the variety of data that it encompasses, big data always brings a number of challenges relating to its volume and complexity. A recent survey says that 80% of the data created in the world are unstructured. One challenge is how these unstructured data can be structured, before we attempt to understand and capture the most important data. Another challenge is how we can store it. Here are the top tools used to store and analyse Big Data. We can categorise them into two (storage and Querying/Analysis).

A. Apache Hadoop

Apache Hadoop is a java based free software framework that can effectively store large amount of data in a cluster. This framework runs in parallel on a cluster and has an ability to allow us to process data across all nodes. Hadoop Distributed File System (HDFS) is the storage system of Hadoop which splits big data and distribute across many nodes in a cluster. This also replicates data in a cluster thus providing high availability.

B. Microsoft HDInsight

It is a Big Data solution from Microsoft powered by Apache Hadoop which is available as a service in the cloud. HDInsight uses Windows Azure Blob storage as the default file system. This also provides high availability with low cost.

C. NoSQL

While the traditional SQL can be effectively used to handle large amount of structured data, we need NoSQL (Not Only SQL) to handle unstructured data. NoSQL databases store unstructured data with no particular schema. Each row can have its own set of column values. NoSQL gives better performance in storing massive amount of data. There are many open-source NoSQL DBs available to analyse big Data.

D. Hive

This is a distributed data management for Hadoop. This supports SQL-like query option Hive SQL (HSQL) to access big data. This can be primarily used for Data mining purpose. This runs on top of Hadoop.

E. Sqoop

This is a tool that connects Hadoop with various relational databases to transfer data. This can be effectively used to transfer structured data to Hadoop or Hive.

F. PolyBase

This works on top of SQL Server 2012 Parallel Data Warehouse (PDW) and is used to access data stored in PDW. PDW is a datawarhousing appliance built for processing any volume of relational data and provides an integration with Hadoop allowing us to access non-relational data as well.

G. Big data in EXCEL

As many people are comfortable in doing analysis in EXCEL, a popular tool from Microsoft, you can also connect data stored in Hadoop using EXCEL 2013. Horton works, which is primarily working in providing Enterprise Apache Hadoop, provides an option to access big data stored in their Hadoop platform using EXCEL 2013. You can use Power View feature of EXCEL 2013 to easily summarise the data. (More information).

Similarly, Microsoft's HD Insight allows us to connect to Big data stored in Azure cloud using a power query option. (More information).

H. Presto

Facebook has developed and recently open-sourced its Query engine (SQL-on-Hadoop) named Presto which is built to handle petabytes of data. Unlike Hive, Presto does not depend on MapReduce technique and can quickly retrieve data.

For the purpose of processing the large amount of data, the big data requires exceptional technologies. The various techniques and technologies have been introduced for manipulating, analyzing and visualizing the big data. There are many solutions to handle the Big Data, but the Hadoop is one of the most widely used technologies

I. HADOOP

Hadoop is a software framework which stores huge amount of data and process it.

Scalable: It can reliably store and process petabytes.
Economical: It distributes the data and processing across clusters of commonly available computers (in thousands).

Efficient: By distributing the data, it can process in parallel **on the nodes where the data is located.**

Reliable: It automatically maintains multiple copies of data and automatically redeploys computing tasks based on failures.

The data can be managed with Hadoop to Distribute the data and duplicates chunk of each data file across several nodes. Locally available resource is used to process, parallel process Handles failover smartly and automatically.

Features of Hadoop

It is optimized to handle massive quantities of various type of data. It Shared Nothing Architecture. Hadoop Replicates data across multiple computers. It provides High throughput with low latency. It complements both OLTP and OLAP.It is **not good** when work is not parallelized. It is **not good** for processing small files because it stores a huge amount of data.

II. HDFS Daemons

Daemons means "Background process".

- Name node
- Data Node
- Secondary name node

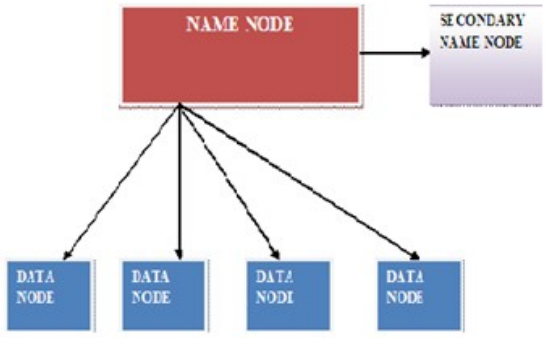

Fig 1: HDFS Daemons

A. HDFS Daemons – Name node(NN)

It is the 'master' machine. It controls all the meta data for the cluster. Eg - what blocks make up a file, and what data nodes those blocks are stored on. HDFS breaks large data into smaller pieces called Blocks. Default block size is 64MB.NN uses RACKID identify. Rack is a collection of data nodes within cluster. NN keeps tracks of blocks of a file as it is placed on various Data nodes. NN manages file related operations such as read, write, create and delete. Its main job is managing the File System Namespace.

B. File System Namespace

File system namespace refers a collection of files in cluster. It includes mapping of blocks to file, file properties and it is stored in a file called FS Image. HDFS supports a traditional hierarchical file organization. A user or an application can create directories and store files inside these directories. The file system namespace hierarchy is similar to most other existing file systems; one can create and remove files, move a file from one directory to another, or rename a file. The Name Node maintains the file system namespace. Any change to the file system namespace or its properties is recorded by the Name Node. An application can specify the number of replicas of a file that should be maintained by HDFS. The number of copies of a file is called the replication factor of that file. This information is stored by the Name Node. HDFS stores multiple data nodes per cluster. It stores each block of HDFS data in a separate file. It performs a Read/Write operation to communicate with Name node and Data node.

III. Map Reduce Programming

Map Reduce Programming is a software frame work. Map Reduce Programming helps you to process massive amounts of data in parallel. It provides a Key- value pair, Job Tracker (master) /Cluster, Task Tracker (slave)/Node, Job Configuration: Application and Job parameters, Interaction between Job tracker and task tracker

Input: Text file
Driver class: Job configuration details
Mapper class: Overrides Map function based on the problem statement
Reducer class: Overrides Reduce function based on the problem statement.

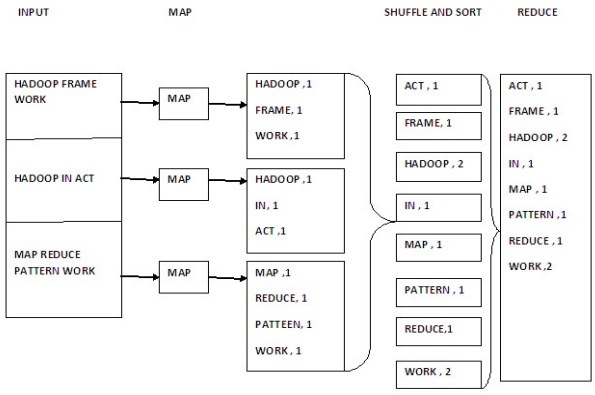

Fig 2 Map Reduce Work

The map task is done by Mapper class and the reduce task is done by reducer class. Input data set is split into multiple pieces of data. The Framework creates several master and slave processes. There are several map tasks works simultaneously. Map workers uses partitioner function to divides the data into regions. Once map is completed reducer work begins. Mapper class tokenizing the given input and after sorts it. In next step reducing process reduces the matching pairs and produces the perfect output.

V. ALGORITHM FOR FILE SECURITY

There are various algorithms available to encrypt and decrypt files. The most frequently used algorithms are Initialization vector and SHA512. The following criteria presents the benefits and limitations of file security.

A. Initialization Vector

Initialization vector (IV) is an arbitrary number that can be used as along with a secret key for data encryption. The use of IV is prevented to repetition in data encryption, making it more difficult for the hacker using dictionary attack to find patterns and break a cipher. If there are repeated sequences in encrypted data, an attacker assumes that the corresponding sequences in the message were also identical. Initialization vector prevents the corresponding duplicate character sequences in the cipher-text.

The ideal IV is a random number that is made known to the destination computer to facilitate decryption of the data when it received. The IV can be agreed in advance, transmitted independently or included as the part of the session setup prior to exchange of the message data. The length of the IV depends on the method of encryption. The IV length is comparable to the length of the encryption key or block of the cipher in use. It is used to encrypt the first portion of the data to be encrypted in file security.

B. SHA 512 ALGORITHMS

To calculate a SHA hash with 512 Bits from sensitive data like passwords, a file is to create a SHA-512 checksum. They additionally provide a shared key to strengthen for the security of the hash. A hash function is an algorithm that transforms hashes an arbitrary set of data, such as a text file, into a single fixed length value the hash. The computed hash value may be used to verify the integrity of copies of the original data without providing any means to derive said original data.SHA-512 is novel hash functions computed with 64-bit words used for secure password hashing. Its use in different shift amounts and additive constants, but its differing only in the number of round structures is virtually identical

The file security on the proposed system shown in fig. 3 encrypts the file for the entered key to generate byte code and encrypted the file. Then the file encryption used the initialization vector and SHA512 hashing key. The flowchart of the encryption process is shown in below in fig. 1. Finally after completing the encryption process the system file guarantee safeness of the keys in HDFS, the system uses with a SHA 512 Algorithm to encrypt the keys. After encrypting the used keys are stored in the HDFS by the authenticated user so that it can be used again in future. Then decryption process as shown in fig 2 to decrypt the file that contains the data again into its meaningful form. The key to access the file in its original form, the system needs to encrypt same key to decrypt the file with its old extension, which is to match with the key and use it's only by authenticated user that's shown in fig. 4

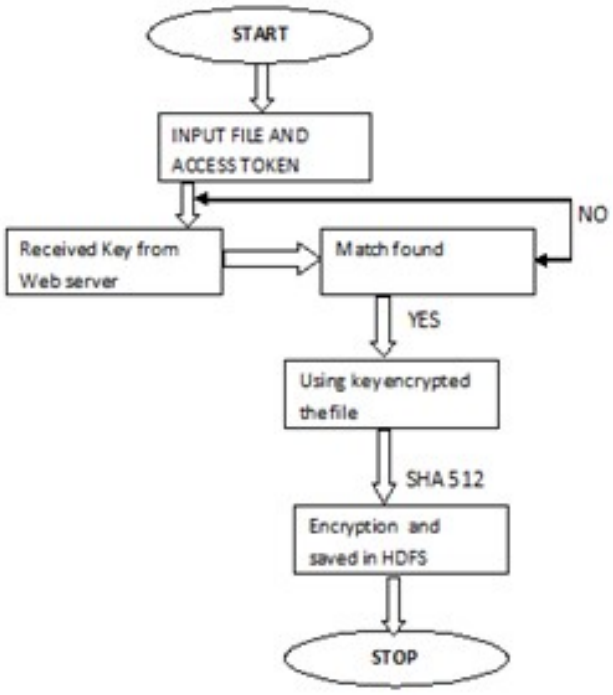

Fig 3 Flowchart for Encryption process

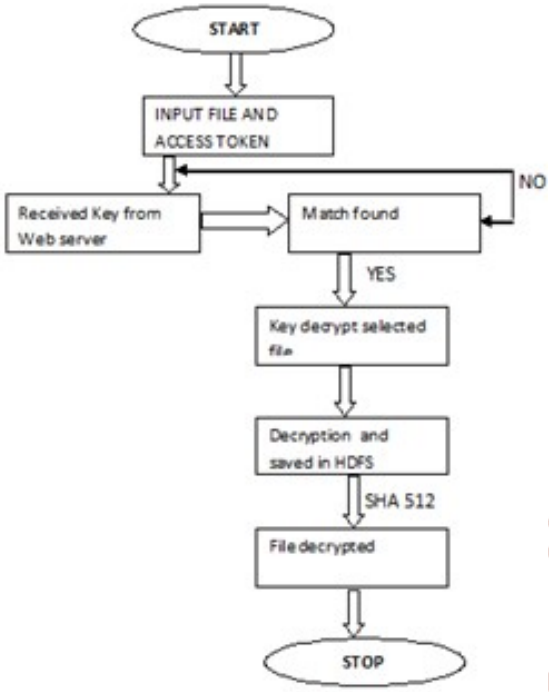

Fig 4 Flowchart for Decryption process

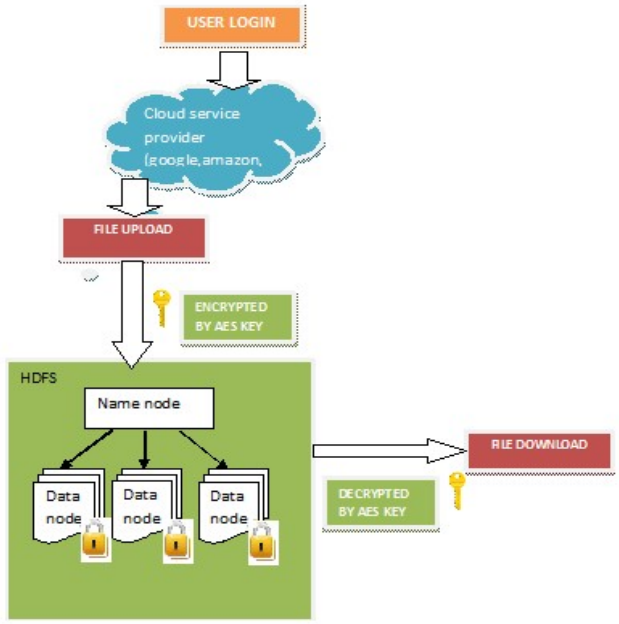

Fig4 Secured File Storage System in Big data with cloud access

VI, PROPOSED WORK

In proposed system, to authenticate user we have used cloud server, which returns unique token for each user who attempts successful login. The token returned by Web server utilized as a part of encryption strategy so it gives information privacy and integrity to the user data.

The files are encrypted before load to HDFS and decrypted when job execution is in progress .The Real Time Encryption Algorithm utilizes the access token as key and Encrypt data (uploaded by user) by SHA 512 with the key [14].

The above problems motivate us to provide a correct, safe and efficient algorithm for securing data saved in cloud storage [14].

The secured file upload and download using encryption and decryption algorithms in the big data tool with cloud access shown in fig 4. Finally after completing the encryption process the system file guarantee safeness of the keys in HDFS, the system uses with a SHA 512 Algorithm to encrypt the keys [16].

After encrypting the used keys are stored in the HDFS by the authenticated user so that it can be used again in future. Then decryption process is to decrypt the file that contains the data again into its meaningful form.

I. Algorithms

Input: Login ID & Password (third party) of client

The following steps are executed

1. Begin
2. Get an access token.
3. User chooses whether to give access to your application
4. User is redirected to your application by Web Server
5. Exchange authorization code for refresh and access tokens.
6. Process response and store tokens
7. Upload the files and the files are encrypted
8. Server verifies credentials & grant access to your application
9. User redirected to your a big data application by server
10. Validation of the client's token
11. Token validation response is processed.
12. Stop

Encryption algorithm

1. Begin
2. Retrieve access token after successful user login
3. Generate key using random key generator
4. Read data from file and encrypt that data with the key, which generated by key generator using SHA 512 Algorithm
5. Append the key to the Encrypted data

6. Write encrypted data in a file and load that file to HDFS
7. Stop

This algorithm suggests the encryption of the files to be uploaded on the cloud. The integrity and confidentiality of the data uploaded by the user is ensured doubly by not only encrypting it but also providing access to the data only on successful authentication. The existed file on the device will be encrypted using SHA 512 algorithm. To enhance security; Random key will be encrypted using SHA 512 algorithm and will be stored in intern server. The authorized user can also download any of the uploaded encrypted files and read it on the system.

Decryption algorithm

1. Begin
2. Retrieve data for decryption
3. Extract key from data
4. Read data from file and implementing SHA 512 algorithm it with the key
5. Pass decrypted data to Map Reduce job submitted by client
6. Combine the output from all working nodes & send it to user
7. Stop

VII. CONCLUSION

This proposed system will provide many operations and flexibility of the data that is been stored in the data base. The user will have the permission for adding, deleting, editing, updating and searching of data. The file is encrypted and then stored in the database and in search operation the file is decrypted and then searched in the database. The user is also provided options to encrypt and decrypt the file by the server. Encryption and Decryption process will be implemented using SHA 512 algorithms for more flexibility.

VIII. REFERENCES

1. Abid Mehmood1, Iynkaran Natgunanathan1, Yong Xiang1, (Senior Member, Ieee), Guang Hua2, (Member, Ieee), And Song Guo3, "Protection of Big Data Privacy" May 9, 2016.
2. Ms. Chetana Girish Gorakh Dr. Kishor M. Dhole2 "A Review On Security Approach In Big Data"
3. G Geethakumari and Agrima Srivatsava "Big Data Analysis for Implementation of Enterprise Data Security " IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 2, No.4, August 2012
4. Kalyani Shirudkar and Dilip Motwani "Big-Data Security" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE) Volume 5, Issue 3, March 2015
5. Khalid Adam Ismail Hammad, Mohammed Adam Ibrahim Fakharaldien, Jasni Mohamed Zain and Mazlina Abdul Majid "Big Data Analysis and Storage" International Conference on Operations Excellence and Service Engineering Orlando, Florida, USA, September 10-11, 2015
6. O. Liu, K.L. Man, W. Chong, and C.O. Chan "Social Network Analysis Using Big Data" O. IMECS 2016, March 16 - 18, 2016
7. Nirali Honest and Atul Patel "A Survey Of Big Data Analytics" International Journal of Information Sciences and Techniques (IJIST) Vol.6, No.1/2, March 2016
8. Prachi Pardeshi, Komal Patil, Priyanka Patil and Komal Chavan4 "A Clustering Based Collaborative and Pattern based Filtering approach for Big Data Application" International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 03,Mar-2016
9. Priyank Jain, Manasi Gyanchandani and Nilay Khare "Big data privacy: a technological perspective and review" Journal of Big Data, 26 November 2016
10. Rishabh Mishra and Dr. Rakesh Sharma "BIG DATA: OPPORTUNITIES AND CHALLENGES" IJCSMC, Vol. 4, Issue. 6, June 2015
11. Samiddha Mukherjee and Ravi Shaw "Big Data – Concepts, Applications, Challenges and Future Scope Information Technology" International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE) Vol. 5, Issue 2, February 2016
12. Shantanu Kalbhor , Hiteshkumar Jain and Kaushiki Upadhyay "Providing Classification And Security Of Big Data In Cloud Computing" International Journal of Technical Research and Applications(IJTRA) Volume 4, Issue 2 (March-April, 2016)
13. Sharifnawaj Y. Inamdar, Ajit H. Jadhav, Rohit B. Desai, Pravin S. Shinde, Indrajeet M. Ghadage,

Amit A. Gaikwad. "Data Security in Hadoop Distributed File System" Volume: 03 Issue: 04, Apr-2016

14. SHASHANK , S. K. Saravanan, G. Rekha "Information Security in Big Data using Encryption and Decryption" International Research Journal of Computer Science (IRJCS) ISSN: 2393-9842 Issue 5, Volume 2 (May 2015)
15. Sophia Yakoubov, Vijay Gadepally, Nabil Schear, Emily Shen, Arkady Yerukhimovich " A Survey of Cryptographic Approaches to Securing Big-Data Analytics in the Cloud"
16. Srinivisan Nagaraj, Kishore Bha midipati, G Apparao, "An Approach to Security Using Rijndael Algorithm", International Journal of Computer Applications, Volume 8– No.5, pp: 365-403, 2010.
17. Venkata Narasimha Inukollu , Sailaja Arsi and Srinivasa Rao Ravuri " Security Issues Associated With Big Data In Cloud Computing", International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.3, May 2014
18. Zaid KARTIT and Mohamed EL MARRAKI "Applying Encryption Algorithm to Enhance Data Security in Cloud Storage" Engineering Letters, 23:4, EL_23_4_06 (Advance online publication: 17 November 2015)
19. Zhang Hongjun , Hao Wenning , He Dengchao and Mao Yuxing, "Survey of Research on Information Security in Big Data" Zhang CSBC 2014