

Arrival Level Data Segregation for Data Management by Emerging Bigdata Analytics and Cloud Plattform

V. Sabapathi* S. Veeramalai* M. Queen Mary Vidya* and P. J. Alin Mercybha*

Abstract : In this paper, arrival level data segregation cabinet for data management by the bigdata and cloud. When user uploads the data in Application in user network interaction, frontend level segregation template should be design for information management and flexible information retrieval .big data has various stream of data processing, challenges and Massive collection of data is available over the network. Data may structure or unstructured but how to retrieving the data it's the very crucial part in the data mining. So we have to make standard cabinet user interface for frontend for information management. This methodology is varying based on App. but frontend information arranges will give a very flexible and efficient way of Information retrieval process. Data cllection, information management and secure level can achieve by this arrival level data segregation protection system .Protection level enhancing can be achieve because of frontend level data arrangement, when hacker try to upload the data (virus) malware in application then the security has been increase via this system of frontend data cabinet design.

1. INTRODUCTION

Bigdata is the vast area in computing pool. Bigdata is development for making unstructured, imperfect data into structural information or actionable information. All the data should be store and making very neat template for structuring information. Bigdata analytics provide analytics of data. Because data is different level and different format is available like text, audio, video. Cloud it's the technology for rends the computing resource tool. Cloud technology for enhancing system resources and making optimum solution over physical resources by virtualization technique. Entry level data segregation is very important for information management. Generallay the data arrangement will made by backend of the application. So the Data segregate or classification of data is important task. Concepts of creating design Graphical User Interface (GUI) pattern has place vital role in Relational Database Management System (RDBMS) [6].

Protection system has been evolve via this system When hacker try to attack such Direct database attack has been reject via this arrangement. Hacker who wants to attack the server such Direct Data attack (DDA) without met webserver, directly crashing the main server is possible when the backend level information management. Generally, Army/Military was not attacker can easily attack. So frontend should be strong and strict then only protection, data collection and information management should be possible in the web application. In backend the data may store in cloud server, cloud system should be highly sensible while protecting the data. Because the Data channel is not protecting, then there is lot of opportunities such as Man in the Middle (MITM) attack like POODLE vulnerability has occur. POODLE as padding Oracle on Downgraded Legacy Encryption has possible via the channel may HTTP/SSL connection. So, secure

* Assistant Professor, Department of Computer Science and Engineering, VEL TECH HIGH TECH Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi, Chennai, Tamilanadu, India Sabapathi2000@gmail.com, veera2000uk@gmail.com, queenmaryvidya@gmail.com, mercybhaalin@gmail.com.

channel such HTTPS/TLS should be used for against POODLE. Data Protection has been achieved via this frontend secure cabinet design as well as secure channel HTTPS/TLS has been used in our proposal s

2. LITERATURE SURVEY

Cloud technology is mainly for minimize the usage of cost of computing resources [1]. Cloud it's the technology for enhancing the system without modifying of physical over the internet. All infrastructure, platform, application softwares are avail via IAAS, PAAS, SAAS. Security is important concern in the cloud system because lots of data vulnerability is in cloud system such POODLE vulnerability in cloud system [5].

POODLE it's the Man In the Middle (MITM). Elliptic Curve Cryptography and Diffie Hellman which provides more efficient for secure key generation and HTTPS/TLS channel using against POODLE [8]. Because POODLE occur in HTTP/SSL channel. Big data which is referring massive data and information maintain, processing, information retrieval such vital concern functioning. Hadoop is technology which has two important subdomain like HDFS (Hadoop Distributed File System (HDFS) and Map Reduce. Hadoop technology is used for Master/Slave structure for cluster of servers maintain across the globe network system. HDFS improves the reliability Hadoop Distributed File System (HDFS) for rapid data transfer achieve through this technology. Map Reduce for parallel clustering technique and sorting the map and executing the file processing. Map reduce will provide fault tolerant enhancing the reliability. The term "Big data [4]" is originated from web search companies for very large distributed data. Cloud computing technology is being used to minimize the usage cost of computing resources. Designing Graphical User Interface for RDBMS has mainly focus of data management [6]. Security enhancing via frontend and backend database from intruders of web application [7]. Data protection such direct database attack has been avoid by this arrival level segregation. This is performed by encrypting the replicas and sending the data dynamically in distributed system over the network [10]

3. METHODOLOGY

A. Data Analytics

Data collection or data entry is vary in each and every human interactions over internet, for example social media like twitter, Facebook is vary if mobile banking, transactions are different. Hence data in banking, business data transactions are very important and more secure is need. So what kind of data to be stored, whether entering/uploading data is structure or unstructured Data has to be analyzing is the very important.

B. Problem Discuss

Bigdata has some problems such Too Much Information (TMI) in while data streaming over network, security issues in data streaming such Man in the middle (MITM) especially data may store in cloud system. Both are very sensitive and take place in information retrieval. Information management which place vital role in every application such social media, banking.

4. PROPOSED SYSTEM

Bigdata, analyzing the kind of information is the very crucial challenge. Data which is entered or uploaded by user is vary in every application. So Every App developer should concentrate UI developing at the frontend. Because entry level data arrangement will make a very efficient way for information management. In this project we propose the STANDARD TEMPLATE CABINET for making structure, semi structure, and unstructured data at frontend. User Interface development takes vital role for developing STANDARD TEMPLATE CABINET SYSTEM as well as in information management. So, arrange of data storing and retrieval can be efficient and flexible.

A. Emerging Cloud

Clouds with bigdata technologies are very essential for current trends. Because emerging of cloud and bigdata ultimate goal is making very optimize data storing and information retrieval in efficient way. For

data can be store in cloud IAAS (Infrastructure as a Service) in virtual storage system. So, sticking these technologies will make very efficient data processing can achieve.

B. Architecture

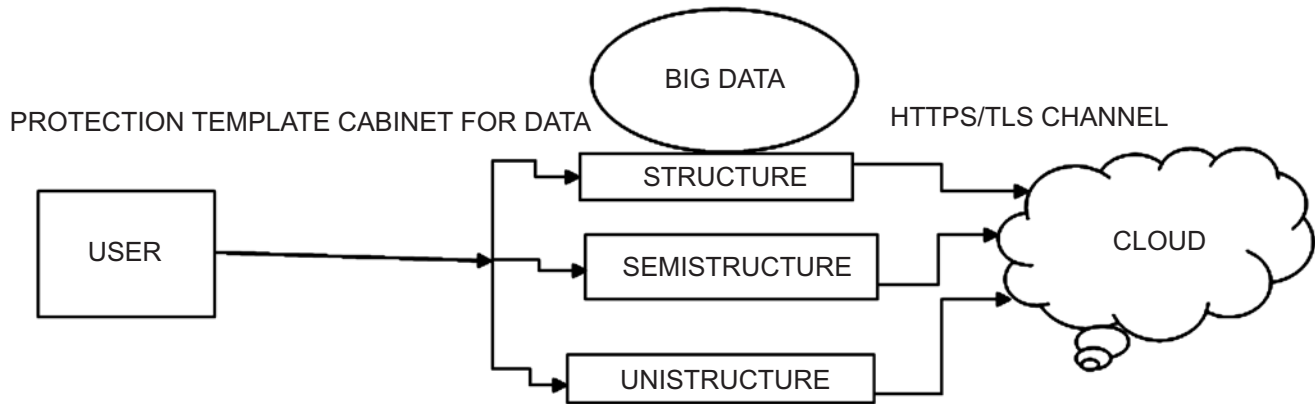


Figure 1: Standard Frontend Cabinet For Information

C. Management

Web application, those may social media, banking transactions, wherever the end user uploading data should be very strong template. Frontend of each application should have strong standard template UI is recommended for data arrangements. Fig[1] shows, where the user enter the information in the application, every app should maintain good template formation so that frontend information management is possible. This Frontend can be vary based on App, because some of transaction of different is differ from social media. Based on App secure level frontend cabinet is design. Data arrangement and stored in cloud server also important task because lot of vulnerability such as POODLE (Padding Oracle on downgraded Legacy Encryption) Man in theMiddle (MITM) will occur. HTTP/SSL base channel has been attacked by POODLE. So Using HTTPS/TLS base secure connection channel has been used in this system. So, backend level security mechanism also important concern. BACKEND of cloud storage system should be more secure using DH and ECC (Elliptic curve cryptography) secure key generation and secure handshaking system [5].

D. Protection Against Direct Database Attack

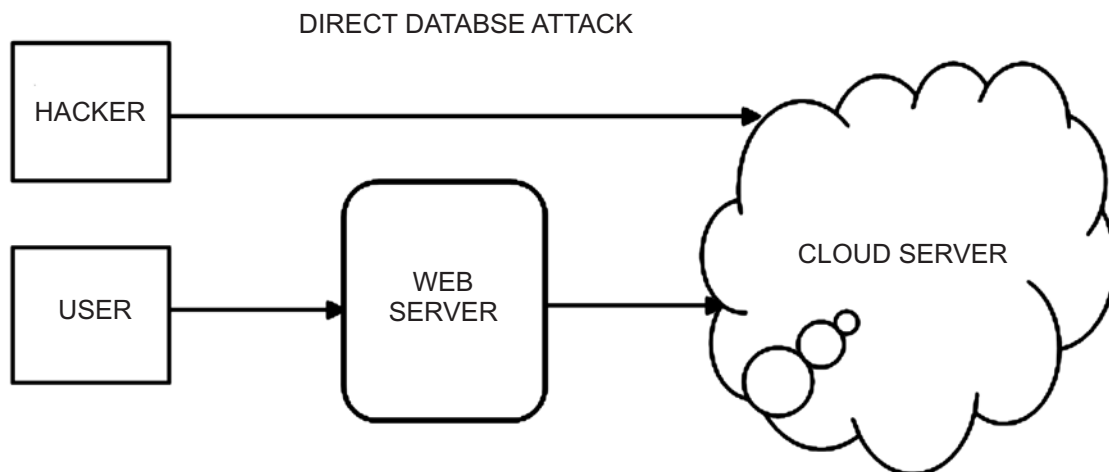


Figure 2: Direct Database Attack (DDA)

Direct database attack has directly uploaded the malware into server system. Hence the backend information management had done in this system. So in our proposed system frontend level data segregation and protection cabinet system has been created.

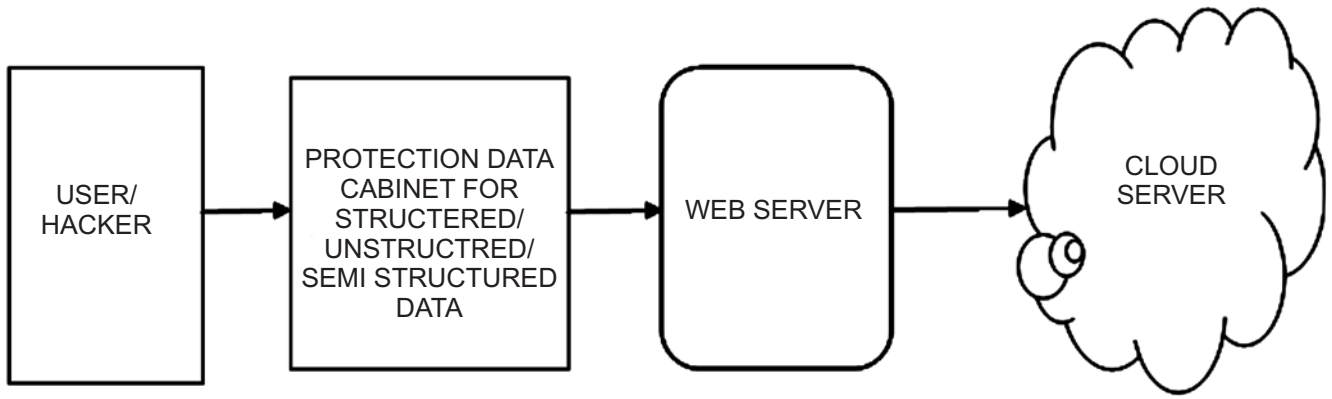


Figure 3: Common Protection Cabinet Against Direct Database Attack (DDA)

So in our proposed system made common standard Protection cabinet for data design for against variety of hackers. Whoever may try to crash the server, but entry level protection will strict, then the all the things will be fine. Information arrangement, data management and protection system can be enhancing this system

E. Structured Data Vs Unstructured Data

80% of business industry data originates in unstructured format, as primary text as stated by Seth Grimes [9]. For example whatsapp, facebook has send /receive message information such date, time has be a structure format. But content may or not be structure. So content, information management is biggest task. Structure data has specific cabinet of data, easiest to parsing data .But unstructured can't be parsing easily.

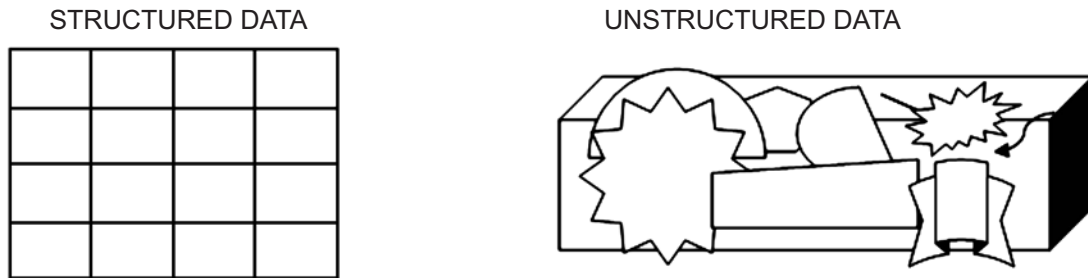


Figure 4: Variety of Data

F. Classification of Data stream

Classification is the process of organizing categories. Using classification algorithms develop the standard template cabinet for categorizing structure, semi structure and unstructured data. Classification algorithms generally called supervised machine learning. Hence naïve bayes algorithm can be used for classification of data. Naive Bayes is a probabilistic classification algorithm based on Bayes. If class variable C and feature variables F_1, F_2, \dots, F_n , the conditional probability (posterior) according to bayes theorem is given as,

$$P(C|F_1, F_2, \dots, F_n) = P(F_1, F_2, \dots, F_n|C)P(C) P(F_1, \dots, F_n)$$

Where $P(C|F_1, F_2, \dots, F_n)$ is the posterior probability, $P(F_1, \dots, F_n|C)$ and $P(C)$ is the prior probability. Naive assumption independence every pair of features given as,

Posterior Probability $[P(H/X)]$, Prior Probability $[P(H)]$ where X is data, H is some hypothesis. According to bayes theorem classification,

$$P(H/X) = P(X/H)P(H)/P(X)$$

$$P(F_1, F_2, \dots, F_n|C) = \prod_{i=1}^n P(F_i|C)$$

In practice, since the evidence $P(F_1, F_2, \dots, F_n)$ is constant for given input and does not depend on the class variable C , only the numerator of the posterior probability is important for classification.

$$P(C|F_1, F_2, \dots, F_n)P(C) = \prod_{i=1}^n P(F_i)$$

With simplification, classification can be done

$$C = \operatorname{argmax}_c P(C) \prod_{i=1}^n P(F_i)$$

5. RESULT AND DISCUSSION

Classification of data which is structured or unstructured at frontend level segregation will provide more efficient for information management. Arrival level data segregation using standard cabinet which give very good knowledge to the user for what type information they are uploaded because the segregation of data can be ensure at the frontend.

6. CONCLUSION

In this paper we define the idea for information management for data processing. Every application should have structured, semi, unstructured template cabinet should provide to the user. Then only user can know the data variety and arrangements of data. This will make efficient and flexible way for information retrieval and make optimizing data processing. Based on App developer can make the data template so that user can upload/entry level arrangement of data is possible at front end. So backend data storage of cloud system has been very easy.

7. FUTURE ENHANCEMENT

In future big data is very high level vital role in data management .Information retrieval is the biggest challenge .So more concentrate on secure design UI is important task.so designing frontend level technical implementation is our future focus.

8. REFERENCE

1. Y. Amanatullah, Ipung H.P., Juliandri A, and Lim C. "Toward cloud computing reference architecture: Cloud service management perspective.". Jakarta: 2013, pp. 1-4, 13-14 Jun. 2013.
2. Hao, Chen, and Ying Qiao. "Research of Cloud Computing based on the Hadoop platform.". Chengdu, China: 2011, pp. 181 – 184, 21-23 Oct 2011.
3. Zhao, Yaxiong , and Jie Wu. "Dache: A data aware caching for big-data applications using the MapReduce framework" INFOCOM, 2013 Proceedings IEEE, Turin, Apr 14-19, 2013, pp. 35 – 39.
4. A, Katal, Wazid M, and Goudar R.H. "Big data: Issues, challenges, tools and Good practices.". Noida: 2013, pp. 404 – 409, 8-10 Aug. 2013.
5. Lu, Huang, Ting-tin Hu, and Hai-shan Chen. "Research on Hadoop Cloud Computing Model and its Applications.". Hangzhou, China: 2012, pp. 59 – 63, 21-24 Oct. 2012.
6. Manoj E. Patil et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,3871-3874 ISSN :0975-9646.
7. Shyam A Gade "Front End and Back End Database Protection from Intruders in Distributed Web Application" Volume 3, Issue 11 (ISSN 2250-2459).
8. Sabapathi,Muniyappan,Danu,Senthil "A study of protocol security-Cloud Architecture" in IIOAB:2016,pp.163-170, Vol 7|9ISSN 0976-3104.
9. M.Giri, Mrudula Sarvabhatla, Chandra Sekhar Vorugunti, "A Secure Mutual Authentication Protocol for Cloud Computing using Secret Sharing and Steganography", IEEE Cloud Computing for Emerging Markets 2014 (CEEM-2014), ISBN: 978-1-4799-6142-9, 15-17 October 2014.
10. Veeramalai Sankaradass*, G.Nalini Priya "A Recommendation System for Dynamic Replication and Securing Data Process", Indian Journal of Science and Technology, Vol 8, Issue 29, article 79163-143787-1, NOV2015
11. M.Giri, Mrudula Sarvabhatla, Chandra Sekhar Vorugunti, "A Robust Ticket-Based Mutual Authentication Scheme for Data Security in Cloud Computing", 2nd IEEE International Conference on Data Science and Engineering ICDSE-2014, ISBN: 978-1-4799-5461-2, 26-28 August 2014.