



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6

Issue: II

Month of publication: February 2018

DOI:

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

Security Perspectives and Challenges in the Deployment of Big Data Technology

Parimala.S

SRM Institute of Science And Technology

Abstract: *In every field the data is one of the most important assets that needs to be handled and performed in an efficient way. The vast amount of growth in the voluminous of data is one of the major issue since it cannot be handled by conventional technical methods. The solution that came as the state-of-the art is the new technology called Big Data. Again it created new problems related to privacy and data security. This paper enlightens some of the security and privacy issues in Big Data.*

Keywords: *Big Data, Data confidentiality, Privacy, Security*

I. INTRODUCTION

Technological advances and novel applications, such as sensors, cyber-physical systems, smart mobile devices, cloud systems, data analytics, social networks, Internet of Things (IoT), smart and connected healthcare, are making possible to capture, process, and share huge amounts of data – referred to as big data – and to extract useful knowledge, such as patterns, from these data and predict trends and Events [1]. Big Data is dynamic and also The big data availability and their usage towards very confidential and sensitive data and privacy data makes the critical situation for the requirement of security and privacy issues.

As everyday data are being collected from applications, networks, social media and other sources Big Data is emerging. Studies have shown that by 2020 the world will have increased 50 times the amount of data it had in 2011, which was currently 1.8 zettabytes or 1.8 trillion gigabytes of data [2]. Most of the application work in real-time systems so these applications must arrange for a suitable network in large volume to create, storage and for processing the large amount of data and information.

II. WHAT IS BIG DATA?

What exactly is big data? A report delivered to the U.S. Congress in August 2012 defines big data as “large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of the information”. [3]. Big data is a term used to refer a large anthology of data sets which is a recent trend in research topics which has many challenges which may include collecting, analyzing and distributes vast amount of data to both public and private domains.[4].

III. DATA SECURITY IN BIG DATA

A. Why Big Data Security Issues Are Surfacing

Big data is now becoming popular even in smaller and medium sized companies due to the factors of cost reduction and easy way to manage the data. The storage in cloud based has coordinated the concepts of data mining and data collection. However the integration of big data and cloud is a very big challenge in integration. The openings may be due that security applications which are designed to store certain amounts of data and information. The management of dynamic data are inefficient by the security technologies as it can manage only the static data. Thus a normal security is not enough to detect the patches in the security for the streaming data. To analyses the big data as well as streaming of data we need Full-time privacy.

B. Secure Computations in Distributed Programming Framework:

The framework of distributed programming makes the parallel computation and utilization of resources and store large amount of data and information.

For instance the Mapreduce framework modularize the input files into multiple chunks of data in the first cycle, and mapper component reads the data in each chunk and manipulate some operations and outputs a key value pairs. In the next part, a reducer component integrates all the value and produces the output. The two major challenges to be faced in this technique are 1) To secure the mappers and to protect the data in case of when the mapper is not reliable.

The Overall MapReduce Word Count Process

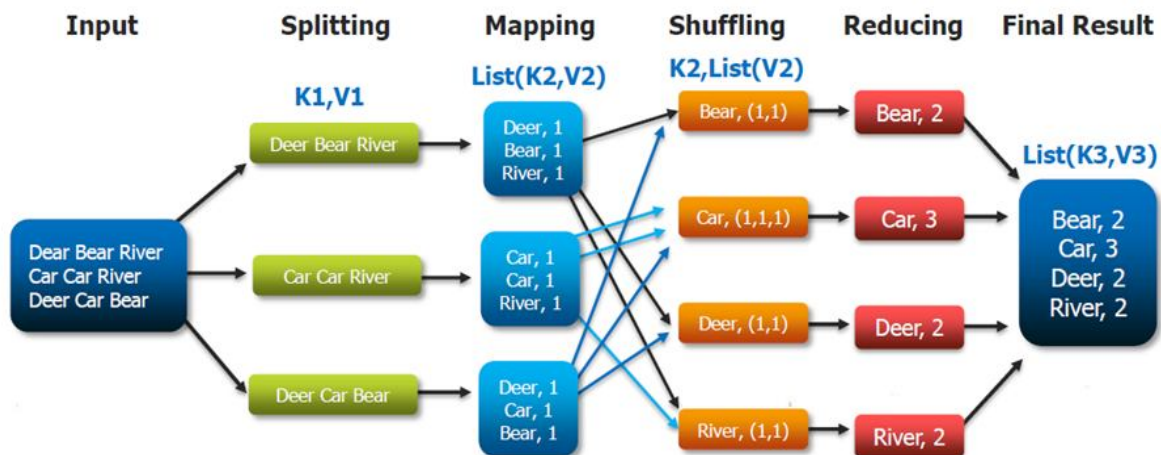


Fig. 1 Map Reduce Architecture

C. Protecting Transaction Logs and Data

When the data is stored in the storage medium, such as logs and any other sensitive information and transactions there will be various levels. For example, the data transfer between these stages gives a vision and understanding about the data to the IT manager.. The organization have to protect these kind of data and should increase the scalability in database storage management. Thus new challenges to be faced for storage management.

D. Security Best Practices For Non-Relational Data Stores

To handle no-relational data one of the popular database is NOSQL. But unfortunately it cannot deal with security and privacy issues and also data analytics. NOSQL creates loopholes for many types of security issues. These dodges include the inability for the encryption of data when it is being generated and stored.

E. Granular Auditing

Consistent auditing of data is necessity whenever the monitoring of data is done. Analysis of various logs created can be done and this information can be done and used to detect all kinds of malicious attacks.

F. End Point Input Validation/Filtering

Data collection is required from many sources to use big data uses cases in organization backgrounds. We may need Security Information and Event Management system (SIEM) which is used to collect the event logs and transactions which was happened from hardware devices and application programs in an y firm network. The validation of input is the major challenge while collecting the data. How far we can trust the data? How we can validate the data and the source from where we obtain the data has any malicious data of input?

G. Real-Time Security/Compliance Monitoring

Real-time monitoring is always been a challenge, the alerts are generated by the Security. These alerts are mostly ignore which leads to false notions. This issue further creates problem and even increase with the volume and velocity of streams of data. One of the main problem is to inspect the real-time data which generated large amount of alerts and also false positives.

H. Cryptographically enforced data centric security

In traditional method the enterprise security focused on the fundamental set-up from malicious attacks. The infrastructure of Big Data is widely spread and also virtualized that it's difficult to self-guard. The encryption of data are learnt by experts to offer protection from the beginning to the end.

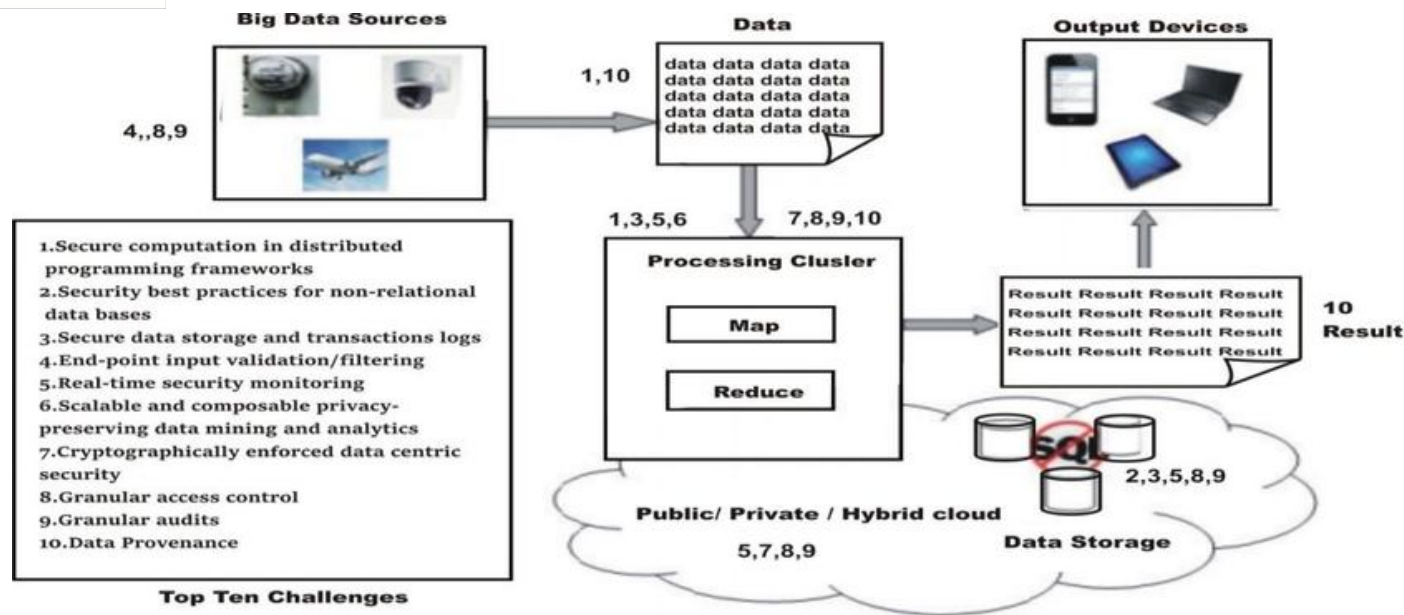


Fig.2 Top 10 challenges in big data technology

I. Granular Access Control

According to Cloud Security Alliance (CSA) there are two essential things in Access Control. The first is to build the access control and to implement a set of strategies that chooses the correct one in the given situation.

J. Data Provenance

The attribution or the derivation of metadata will grow in intricacy due to large provenance graphs which are generated from provenance programming which are enabled in the programming environment of big data application. To detect the metadata dependencies for security or privacy applications in intensive computationally.

IV. CONCLUSION

This paper is about the study of assorted security and privacy issues and challenges in the Big data Technologies. Even Big data is a very challenging domain for research, through some better analysis of the massive amount of data, there is an impending capability for making the better advancements in many scientific areas and also improve the profit and also the success for many organizations. Moreover these challenges will need some transformation of solutions and will not be addressed before this latency is realized. Sustenance and encouragement for the fundamental research is need for speaking out these technical challenges if we want to achieve the potential benefits that are being promised by the Big Data Technologies

REFERENCES

- [1] E. Bertino, Big Data – Opportunities and Challenges, Panel Position Paper, Proceedings of COMPSAC 2013
- [2] Mearin, Lucas. "World's data will grow by 50x in next decade, IDS study predicts". Computerworld.com, Web Accessed on May 2016.
- [3] IHTT: Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry. 2013, IHTT: Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry. 2013, <http://ihealthtran.com/wordpress/2013/03/iht/C2%B2-releases-big-data-research-report-download-today>
- [4] Parimala.S, Dr MV Sreenath, "Improving the effectiveness of e-learning through the behavior of proactive e-learners using big data concepts "Scientific Transactions in Environment and Technovation April 2016, Vol.9 Issue 3 pp-139-143
- [5] E. Bertino, R. Sandhu, Database Security – Concepts, Approaches, and Challenges, IEEE Transactions on Dependable and Security Computing 2(1): 2-19, January-March 2005



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)