Opinion
Volume 11:8, 2022

Journal of Industrial Engineering and Management

ISSN: 2169-0316 Open Access

Static Data Responsible for Encryption Management Module

Mike Stelor*

Department of Mechanical Engineering, University of Dudley, Dudley, UK

Introduction

Data that is large in size (probably petabytes, exabytes, or zettabytes) and generated by numerous social media platforms and businesses is referred to as big data. Consequently, utilizing other databases software makes it quite challenging to store, secure, and process such a large volume of data. Also, there are a lot of problems, like how to manage huge amounts of structured, unstructured, or semi-structured data, where to store such a large amount of data, how to process it, and security problems, which are really the biggest problem because large amounts of data can be hacked or attacked. Hadoop, which stands for Highly Archived Distributed Object-Oriented Programming, is a platform that uses multiple nodes to store, manage, access, process, analyze, and distribute big data. It is a distributed, open-source framework technology that receives data from big data clouds in a variety of formats. It has one expert hub having metadata of client hubs and many slaves/client hubs having real appropriated information. Two parts make up Hadoop: Data is stored on the Hadoop Distributed File System (HDFS), and data is analysed on Map Reduce. In a Hadoop environment, HDFS is more scalable, distributed, and reliable.

Discussion

Before securing data in motion, the first step is to secure data at rest, which refers to data that is stored on a specific medium. Some proposed techniques include encryption, anonymization, data masking, and data erasure. It is important to create an environment in which data can safely move or transfer from one location to another. Hadoop, which stores and processes a large amount of data, provides different mechanisms while data is in motion.

Token of block access (BAT): Brendan Eich, the JavaScript programming language's creator and Mozilla and Firefox's co-founder, proposed BAT. Occasionally, clients access the data directly from the data node after accessing the data block ID from the Name Node. As a result, BAT is utilized in this instance to guarantee that authorized users can only access data blocks stored in data nodes. To determine which data node contains the required data blocks, any user who wishes to access any data node first accesses the name node. The taken BAT is then issued by the name node, and the data node uses it to identify authorized clients. After that, Data Blocks validates the token and issues a Name Token, granting the NameNode permission to securely access its data blocks. Hash-256 Algorithm was designed by National Security Agency (NSA). This mechanism is implemented b/w user and name node and ensure authentication and manages data nodes. The user authenticates himself to NameNode by sending a hash function. First user sends hash function to get access of data. name node also produces the hash function. After this both hash function compares, if both are correct then data can be accessed successfully. Oracle, Microsoft, and IBM all make use of this technology to encrypt database files. TDE is included in Microsoft SQL Server 2008, 2008 R2, 2012, 2014, and 2016 by Microsoft. It is a HDFS abstraction that is used for transparent encryption. a special directory whose data will be transparently decrypted for read and encrypted for write. The key for the encryption zone is generated when it is formed. Each record in encryption zone is determined by its own exceptional Information Encryption Key (DEK). Only an Encrypted Data Encryption Key (EDEK) can be handled by HDFS. After this EDEK has been decrypted, clients can use DEK for read/write operations. The Hadoop Key Management Server (KMS) is in charge of managing DEK and other encrypted keys in the Hadoop framework.

As it helps in encrypting data from hackers, same as it has some limitations. It is not suitable for data processing. Whenever map reduce task run on encryption zone data, complete data file is first decrypted for processing and after running again encrypting the data. So, a lot of time is wasting in this process which badly affects its performance. Abnormal Directory (AD), Abnormal user (AU), Abnormal operation (AO) and Block proportion (BP). Hadoop have fixed directory to save files if any directory is out of range then it means that there happened an attack. And then it finds the suspicious block that contain abnormal directory and calculate suspicious block proportion.it also detect the Abnormal user. If any of these four dimensions gives warning, it means there happened an attack. And then investigator's start to investigate the problem according to warning message. In Application layer investigation based on Hadoop logs i.e. Name Node, data node. Management of user information and Kerberos authentication. The purpose of the user information management sub module is to manage the user's information authentication in the big data platform. The Kerberos authentication management sub module manages the services of Kerberos authentication in the big data platform. Its functions include Kerberos service startup and shutdown; Kerberos service parameter configuration, Kerberos authentication key management, and Kerberos ticket management. Users can be created, their information viewed, deleted, and their authentication modified through the use of user information management.

Even though each approach does their best to reduce security risks and make Hadoop a secure framework, we cannot say that all security risks have been eliminated. Cybercrime is on the rise, and in order to combat it, we must implement the best security measures possible. The ChaCha20 algorithm is our recommendation because it offers three-dimensional security, improved speed, memory efficiency, is simple to implement, and is flexible, but it does not guarantee authenticity because all approaches are designed to address a specific challenge. We recommend the Kerberos method for authentication because it guarantees user authenticity. Combining these methods will provide greater assurance from a security standpoint. This paper outlines the Hadoop platform's security issues and suggests effective solutions. In order to keep Hadoop a secure framework for storing, accessing, and processing large volumes of big data, various security issues are discussed as well as solutions to them. The discussed strategies are implemented in the HDFS layer to ensure its security and dependability while maintaining performance standards. Mixes of these approaches can likewise be utilized. While each of these methods improves security to a great extent, there are some drawbacks that cannot be overlooked [1-5].

Conclusion

Beyond limiting access to the key management server, you should also limit access to the keys themselves based on user and group. The users and group access can be defined on a system level, or at the level of each key.

***Address for Correspondence:** Mike Stelor, Department of Mechanical Engineering, University of Dudley, Dudley, UK, E-mail: stelor375@edu.in*

Received: 01 August, 2022, Manuscript No. iem-23-85366; **Editor assigned:** 03 August, 2022, Pre QC No. P-85366; **Reviewed:** 16 August, 2022, QC No. Q-85366; **Revised:** 22 August, 2022, Manuscript No. R-85366; **Published:** 29 August, 2022, DOI: 10.37421/2169-0316.2022.11.164

When you create a key you can define the restrictions on user and group access. As an example: There is an AES encryption key available on the key management server used to protect an employee's personal data. It is restricted so that only members of the Human Resources group can use that key. So any individual with "Human Resources" defined as their individual or group role can successfully request that key, all others are turned away.

References

1. Nambiar, Shruti and John TW Yeow. "Polymer-composite materials for radiation protection." *Appl Mater Interfaces* 4 (2012): 5717-5726.
2. Liu, Fan and Xiaohong Wang. "Synthetic polymers for organ 3D printing." *Polymers* 12 (2020): 1765.
3. Eghtedari, Yas, Lawrence J. Oh, Nick Di Girolamo and Stephanie L. Watson. "The role of topical N-acetylcysteine in ocular therapeutics." *Surv Ophthalmol* 67(2022): 608-622.
4. Fu, Weiqiong, Hanxiao Zhang and Fu Huang. “Internet-based supply chain financing-oriented risk assessment using BP neural network and SVM.” *Plos one* 17(2022).
5. Haldane, Andrew G. and Robert M. May. “Systemic risk in banking ecosystems.” *Nature* (2011): 351-355.

How to cite this article: Stelor, Mike. “Static Data Responsible for Encryption Management Module.” J Ind Eng Manag 11 (2022): 164.