

CP70042E & CP7HA42E – Network and Systems Security Written Assignment

Table of Contents

Introduction.....	3
Big data and its network security analysis.....	4
Unique or new access control challenges for securing big data.....	4
Existing work approaches for securing big data.....	5
Conclusion.....	8
References.....	9

Introduction

Big Data describes the enormous quantities of information being generated everywhere in the world. Some or all of Big Data may be in an unstructured format. It seems to reason that in today's data-driven culture, many patterns and correlations would be presented within the enormous amount of data. The modern economy places a premium on data, making it crucial to transform raw data into information that can be used to generate business intelligence and, eventually, grow a company's operations. More data equals more precise information, which in turn implies better decision-making and more successful ways of planning, regardless of the size, location, market share, customer segmentation, etc. of the organisation. When data is organised in digestible chunks, it can be put to good use by businesses looking to grow by learning more about their customers' wants and needs, the kinds of products they should make and the quality of service they should provide, as well as ways to save money without sacrificing quality and stimulate the economy. So, in general, Big Data provides enormous benefits for companies.

Big data and its network security analysis

The term "Big Data" is used to describe a dataset that is not just massive in size but also likely to grow rapidly in the future. The typical suite of data management technologies is inadequate for this massive and intricate data set. Big data, put another way, is just a lot of data. As a result of improved marketing efficiency and increased customer satisfaction, a company's bottom line might increase thanks to the incorporation of big data into its internal processes. Businesses that effectively leverage it get an advantage over their competitors who do not since the former are better able to make timely and informed strategic decisions. For instance, by illuminating hitherto unseen elements of a company's target demographic, big data may aid in the improvement of marketing, advertising, and promotional efforts (Angin et al, 2019). Companies may better meet the needs of their consumers, both now and in the future, with the use of data analysis.

An evaluation of network security looks for vulnerabilities that could be exploited to steal information or halt production. The purpose of a network security audit is to locate any weak spots that might be exploited by hackers to gain unauthorised entrance to your system and compromise sensitive data. When analysing network security data, security decision-makers seldom collaborated with their peers from other firms in the past. This issue is complicated by a number of variables, including the need for secrecy, a lack of personnel, poor training, and outdated tools. As a result, the go-it-alone approach has major implications due to the fact that analysing network security typically requires a massive human inspection and analysis of abnormal data (Zhang, 2018). Since most organisations use proprietary software or customised versions of off-the-shelf tools to seek for dangers, manual comparisons are necessary when comparing observations with reports from other enterprises. As a result, many hours are wasted when members of different teams idle away their time waiting for their co-workers to respond to emails, read them, examine them, and hopefully take action on them while also balancing their own workloads.

Unique or new access control challenges for securing big data

When it comes to protecting massive volumes of data, there are likely to be a great deal of obstacles in the way. These problems are widespread and may be found in all aspects of the big data and cloud computing infrastructures respectively. The storage of significant volumes of data on the cloud requires an extremely high level of vigilance. In order to meet the

problems that have been presented, we need to design reliable techniques for security. To list just a few of the many concerns that need to be addressed, the following are some of them:

- It is abundantly clear that the future of data analysis will revolve around the utilisation of formats such as these. This is because recent advancements in computing techniques for raw data and non-relational databases have made it abundantly clear that the future of data analysis will depend on the utilisation of formats such as these (NoSQL). There is a chance that it will be difficult to provide absolute assurance that these innovative practises and technology do not come with any inherent dangers (Hossain et al, 2019).
- Technicians that specialise in the eradication of malware need to be able to clean up after themselves in a timely manner. In order to stop any additional harm, antivirus and anti-malware programmes need to keep a close eye out for malware and find it as quickly as possible.
- A number of different encryption programmes are utilised during the authorisation process, as well as the storage phase, in order to ensure that the secrecy of the data is preserved at all times. When these encryptions are used on data that was created on a number of different systems, there is a definite risk that they will not offer the same results. This is a possibility that cannot be ruled out (Amanullah et al, 2020).
- Because of the magnitude of the data, conducting security audits on data sets that span many petabytes may be a challenging endeavour. The great majority of big data platforms are constructed using a cluster design. As a result, the individual nodes and servers that comprise these platforms are susceptible to a broad variety of security issues.
- Big Data system administrators are the only ones who have the sole right to engage in data mining. This privilege does not extend to anybody else. Users should be notified if they have not been granted authorization to access the system and the application should keep a record of who has signed in and when they did so. In addition, the application should record who has signed in it (Wazid et al, 2019).
- In the event that the environment for big data is not kept up to date and maintained, it is likely that data may be destroyed.

Existing work approaches for securing big data

Safeguarding Distributed Programming Framework: In order to ensure that the distributed system is completely risk-free, it is necessary to put into action procedures such as Kerberos Authentication and to adhere to widely accepted standards of good security practice. Personally Identifiable Information (PII) needs to be kept in a separate location from other kinds of data in order to protect the confidentiality of the information. File access controls and predetermined rules with required access control can be established in order to prevent untrusted programmes from causing sensitive data to be released from devices. These measures can also be taken to prevent unauthorised access to the files themselves (Othman et al, 2018).

Securing Non-Relational Database: Attacks using the NoSQL injection technique are more prevalent in non-relational databases than they are in relational databases. Because we hash passwords before encrypting them, they are protected from the attacks described above. There is a wide variety of algorithms that can be utilized, such as the end-to-end encryption standard (AES), RSA, and Safe Hash Algorithm 2. Security is a significant problem in NoSQL databases because authentication and encryption are either not supported at all or do not work as intended in these systems. The following is a list of some of the risks that are associated with using NoSQL databases: In the vast majority of configurations, administrative access or authentication will be rendered inoperable. The system that stores passwords is extremely vulnerable to security breaches (An et al, 2019). When NoSQL databases were first made available, they did not come equipped with the security and privacy functions that were intended for them. Examples of these characteristics include encryption, authenticating users, and authorising access to resources. The fact that the sharding process is performed in a decentralised manner raises additional concerns regarding data safety.

Transaction Logs and secure Data Storage: Big data can only be trusted if it is stored securely. If an unknown user has acquired unauthorised access to a file and made changes to it, you can find them with the help of Secure Untrusted Data Repository (SUNDR). A cryptographic identifier, like a signed message digest, can be used to uniquely identify files. The features required for storing massive amounts of data are often provided by scale-out or clustered NAS systems. This shared storage solution uses parallel file systems spread across multiple storage nodes to scale out to meet capacity or increasing computation requirements

and to manage billions of files without the performance deterioration that comes with standard file systems as they grow (Bhatia and Sood, 2018).

Big Data Cryptography: By creating scan and filter techniques, like the searchable symmetric encryption (SSE) protocol, businesses may enable Boolean queries on encrypted data. Without the capacity to send and receive data, it would be difficult for anybody or any organisation to function in the modern world. The daily data handled by a company's systems is crucial to the company's ability to function. The potential for sensitive information to be included in big data transfers is a security concern. Recent developments in information technology have opened up new opportunities for economic, personal, social, educational, and scientific pursuits for everyone. Cryptography, sometimes known as the study of "Secret Writing," is a way to ensure the confidentiality of communications across an unreliable medium. Information may be encrypted using an algorithm and a key to keep it secure. Unless you have the right key that was used to encrypt the data, you won't be able to read it. One-way and two-way encryption are only two examples of the many types of cryptography that serve various needs. By encrypting the key only in one direction, we increase the security of the actual encryption and decryption process (Khan and Ansari, 2019). This key ought to be used exclusively for protecting and revealing secret information.

Security monitoring and Real-Time compliance: Ticketing and tokenization are used to make it possible for nodes on an untrusted network to communicate with one another in order to exchange data. Log monitoring and firewalls for servers and routers are two examples of simple security measures that can be implemented. In addition, we have the responsibility of protecting not only the cloud itself but also the underlying infrastructure and any applications that are operating on it. Near real-time data collection, analysis, alerting, and event tracking are the four main components of real-time security monitoring (Angin et al, 2019). These components are performed for all critical devices within a company's network. Security monitoring is the process of gathering and analysing data to detect malicious activity or unauthorised system changes on your network. It is also the process of determining which behaviours should trigger alerts. Security monitoring is also known as security information monitoring (SIM) and security event monitoring (SEM).

Access Control: User access is defined by access control, which determines whether or not a user is permitted access. The most important thing to do is to devise and implement a set of procedures that can reliably determine the best course of action. Access control is a method of

computer security that determines who or what can view or make use of a system's resources within a computing environment. It was a fundamental principle of security theory that contributes to the safety of businesses (Zhang, 2018). When it comes to protecting one's belongings, one has the option of utilising either physical locks or a more complex logical system.

Conclusion

The report concluded that there are several storage facilities for files and transactional information. When information is transmitted manually between levels, the IT manager has full control over what data is shared and when it is transferred. Unfortunately, as data sets expand in size, the significance of scalability and availability grows in tandem with it, necessitating the use of auto-tiering for big data storage management. Because of the lack of visibility into the data's physical location, auto-tiering methods increase the storage system's susceptibility to attack. This will prevent sensitive information from being disclosed to unauthorised parties. If we are serious about preventing the disclosure of sensitive information, it is enough for us to perform the standard security checks on the system. It is the responsibility of administrators working in the cloud or in a virtual environment to scan for all nodes and mappers and look for indications of manipulated data.

References

- Angin, P., Bhargava, B. and Ranchal, R., 2019. Big data analytics for cyber security. *Security and Communication Networks*, 2019.
- Hossain, E., Khan, I., Un-Noor, F., Sikander, S.S. and Sunny, M.S.H., 2019. Application of big data and machine learning in smart grid, and associated security concerns: A review. *Ieee Access*, 7, pp.13960-13988.
- Amanullah, M.A., Habeeb, R.A.A., Nasaruddin, F.H., Gani, A., Ahmed, E., Nainar, A.S.M., Akim, N.M. and Imran, M., 2020. Deep learning and big data technologies for IoT security. *Computer Communications*, 151, pp.495-517.
- Wazid, M., Das, A.K., Hussain, R., Succi, G. and Rodrigues, J.J., 2019. Authentication in cloud-driven IoT-based big data environment: Survey and outlook. *Journal of systems architecture*, 97, pp.185-196.
- Othman, S.M., Ba-Alwi, F.M., Alsohybe, N.T. and Al-Hashida, A.Y., 2018. Intrusion detection model using machine learning algorithm on Big Data environment. *Journal of big data*, 5(1), pp.1-12.
- An, S., Eom, T., Park, J.S., Hong, J.B., Nhlabatsi, A., Fetais, N., Khan, K.M. and Kim, D.S., 2019, August. Cloudsafe: A tool for an automated security analysis for cloud computing. In *2019 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)* (pp. 602-609). IEEE.
- Khan, M. and Ansari, M.D., 2019. Security and privacy issue of big data over the cloud computing: a comprehensive analysis. *IJRTE-Scopus Indexed*, 7(6s), pp.413-417.
- Zhang, D., 2018, October. Big data security and privacy protection. In *8th international conference on management and computer science (ICMCS 2018)* (Vol. 77, pp. 275-278). Atlantis Press.
- Bhatia, R. and Sood, M., 2018, December. Security of big data: a review. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 182-186). IEEE.