DALHOUSIE UNIVERSITY

CSCI 5408 – Data Management, Warehousing, Analytics

Assignment 3

Work done by,

Name: Guturu Rama Mohan Vishnu

Banner ID: B00871849

Email: rm286720@dal.ca

DECLARATION

I, Guturu Rama Mohan Vishnu, declare that in assignment 3 of CSCI 5408 course, summarizing the paper is not done programmatically or using any online or offline tools. However, the webpages or the domain mentioned in this document are visited manually, and some useful information is gathered for education purpose only. Information, such as email, personal contact numbers, or names of people are not extracted. The course instructor or the Faculty of Computer Science cannot be held responsible for any misuse of the extracted data.

Problem #1: Database Administration and Security Reading

Summary:

Day-by-day, the release of data in this world gets skyrocketing. But when this much amount of data is being released, it also has to be stored somewhere for any purpose possible. That's when the NoSQL database comes into the picture. NoSQL databases will exhibit the feature of storage and data retrieval without us defining a schema before creating it. It is one of the reasons why most companies use it. But the drawback with NoSQL is the highly important security and privacy issues. In this research paper, we are going to study the issues regarding the security of the data and it also lists the solutions to protect the data.

To begin with, an SQL database is a traditional structure of collecting tables that store a specific set of structured data. The reason why NoSQL (Not only SQL) databases were introduced is the fact that the traditional systems could no longer handle the amount of data generated. NoSQL has the power over other relational databases because of its schema-less structure, fast response, handling of semi and unstructured data. It also can do more than just perform a query and throw an output. But the flaw here is that these databases don't come with a data encryption layer. The absence of encryption support and not sufficient correspondence between the client and the server, are the things that are prompting to become a genuine danger for the people who use it.

To address this issue of protection and security, beneath are the proposed solutions in this paper:

(a) Pseudonyms-based Communication Network

This solution proposes to utilize two imperative protocols, the RSA (Rivest-Shamir-Adleman) and the Diffie Hellman. The essential thought behind this arrangement is to keep the client's personality unknown by allowing them to approach different administrations on the information base by checking in just a single time and not every time. This one time is the point at which they get associated with the framework at first, and the exchanges completed by the clients can't be connected to them since their character is covered up. This arrangement helps in safeguarding the personality of the clients. The structure and operations expand Brand's credential system, while it comprises of four gatherings: the users U, a central Identity Provider as IP, the Service Providers

SPs, and the organization for giving and approving certifications. Users are entities that get accreditations and are known to Service Providers just through their pseudonyms. Also, the Credential Authority CA forestalls the sharing of accreditations or aliases and ensures that clients who enter the framework have a public and mystery key that makes them one of a kind to the framework. To be brief, it is beyond the realm of possibilities for various clients to work together and show a portion of their qualifications in a Service Provider as well as to get a certification for a client that they couldn't acquire.

(b) Monitoring, Filtering, and Blocking

NoSQL databases can't identify and afterward impair malignantly jobs and queries. In a cloud environment, no data concerning the correspondence of nodes in the cluster or client association details or information modifying data (in any event, altering or erasing) is recorded. By and large, since there are no log records, a difficult issue is to distinguish occurrences of an information break or pernicious information misfortune in the cluster. Continuous security systems exist in big data innovations, bringing about rapid information investigation. We can settle this through Kerberos. The underlying validation of the client should be possible through Kerberos, and next-level authentication should be possible for getting to MapReduce.

In this paper, we have talked about significant security concerns with respect to NoSQL databases. In NoSQL databases, Kerberos is utilized to confirm the clients and information nodes. In particular, to guarantee fine-grained approval, information is gathered by their security level. Different procedures for moderating the attacks on NoSQL databases have additionally been talked about, alongside the proposed security and protection arrangements of NoSQL databases.

Scope of Improvement:

- Management of the encryption keys is likewise similarly significant which has a scope to get improved
- Isolation of jobs among administrator and staff additionally helps in diminishing information breach which goes under filtration.

References:

[1] G. Vonitsanos, E. Dritsas, A. Kanavos, P. Mylonas and S. Sioutas, "Security and Privacy Solutions associated with NoSQL Data Stores," 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA, 2020, pp. 1-5, doi: 10.1109/SMAP49528.2020.9248442.