Database Application- Assignment Report

INTRODUCTION

This research report is going to demonstrate the advantages and disadvantages of two popular databases that are one of the leading pioneers in the industry, Oracle Database and MongoDB. These two will be evaluated in terms of different aspects such as their resources requirements, scalability, data integrity and complexity. These findings are supported by Oracle and MongoDB documentation, reading articles and experience gained from assignments handed out this semester at RMIT.

Overall, Oracle is a relational database management system (RDBMS), which means data is stored in a systematical format using columns and rows. On the other hand, MongoDB is a NoSQL document-oriented database storing data in BSON format (binary serialized JSON format) and grouped in collections (a counterpart of tables).

ORACLE

About Oracle, it has a long history with great innovations for database management tools since the late of 70s. Data in Oracle is stored in a structured format (tables). Therefore, data retrieval of a specific value is easy to locate and access. Oracle is called relational database because their tables may or may not relate to others. Due to this nature, data needs to be predefined in database schemas to remain high referential integrity. Referential integrity establishes a set of rules on columns and data type, so that only if the data manipulation for a specific column which matches the value in a column of a related table is allowed (CRUD operation). This integrity is enforced with constraints such as primary key, foreign key, unique, not null, etc. This increases the consistency within Oracle database. As a result, advanced queries can be run across multiple tables at once. Nevertheless, designing robust schemas and build them could be difficult.

Because of the high-level database design, forming good query requires a good and critical understanding of the database. In Oracle, queries are written in structured statements, also called SQL (Structured Query Language). For example, SELECT … FROM … WHERE … Oracle also provides artifacts and tools that you can use to improve your query for data manipulation such as materialized view, joining tables, execution plans and splitting query. Data can be also standardized to get better result from the query, because some different data may say the same thing within the database. Queries in Oracle are not only for CRUD operations but also can perform some other different tasks, such as creating trigger, calculating and summarizing data. In addition, Oracle allows transaction grouping, which let user have a group of data transactions (low-level tasks) divided into smaller transactions and executed independently for data retrieval or modification. And data can be combined during fetching in a reliable way. This utilises the performance in as least time as possible.

About resource requirements for Oracle, it varies between the size of the business or environment. For example, a small environment for Infrastructure Management Oracle

database server requires 16GB of RAM, 170GB for the database and 15000 RPM drive or a tier 1 SAN storage (2-4 GBps SAN dedicated channel), according to BMC Software 2016. Oracle provides very informative documentation for resource requirements for different

types of needs and demands on their website.

With the security aspects, Oracle provides to ability to deny privileges in groups and roles differentiate from one another as well as external authentication such as LDAP and Active Directory. Network encryption and the message digest are support as well. In terms of problem, Oracle database may experience SQL Injection attack. This is when the attacker tries to pass a SQL string to the database for malicious activity.

Regarding scalability, RDBMS focuses on ACID properties, which may come down to consistency. Every modification you make to the database could affect the other to remain consistency. If tables spread across several servers, it could be very challenging to scale. Therefore, relational databases are designed to run on single servers to maintain integrity and tend to scale up on single machines. This could put a high cost on infrastructure and hardware.

MONGODB

Data in MongoDB are divided into Databases, Collections and Documents. Documents. Documents are presented in a set of JSON-like key-value pairs and applied in dynamic schemas. Schema-less means that users can have any data type in a separate document. Thanks to this, MongoDB is very versatile with large and unstructured data with variety. This also makes MongoDB lack of referential integrity.

Queries are performed by using commands such as insertOne, updateOne, find(), count(), etc. MongoDB also has indexing and is equipped with analyzing query (with $hint and $explain) to help user maximise the efficiency of a query. Due to its simplicity MongoDB tends to have fewer tools and artifacts for users to tinker with. Also due to this, it greatly improves the speed and is estimated to be 100 times faster than relational database, according to Data Flair 2018. One of the disadvantages of MongoDB is that it does not support joins that are very crucial in relational database like Oracle. Instead, users need to aggregate data by $lookup and aggregation pipeline. Not until the release of MongoDB version 4.0 that the users receive multi-document ACID transaction support.

According to FotoWare 2019, MongoDB requires around 1GB of Ram for every 100.000 assets. In regards to disk space, each asset requires 10Kb and is kept approximately 10 days in MongoDB if deleted. One finding states that with the default MongoDB configuration, it could use up to a half of 1 GB for its cache size (Rowe 2019). This high-memory usage is one of MongoDB weak points.

In terms of security issues, MongoDB is highly subjected to ransomware attacks with its default security configuration, where data is deleted from original and the victim is asked to pay an amount of cryptocurrency to obtain back to data (Claburn 2017). MongoDB also has

weak authentication, authorization, encryption by default and may require developers to configure and test independent (Kirkpatrick 2013). Nevertheless, classified as a NoSQL database, MongoDB is unaffected to SQL Injection like RDBMS is.

Scalability with MongoDB fits the ongoing demands of businesses when it comes to upgrading the database system. Regarding throughput and data volume, MongoDB supports the users with horizontal scaling with sharding among serval servers connecting to the same application, which helps minimizing downtime and increasing capacity and availability to manipulate the data. In other words, one database could be partitioned into smaller ones. With this approach, the overall cost of hardware could be deducted; however, this could result in more complex infrastructure and maintenance.

CONCLUSION

Overall, each database has its own advantages and disadvantages. Users should take into consideration on their needs, requirements, budget and goals when choosing one. For example, if your data is unstructured or complex or pre-defined schemas is not required, MongoDB could be a good choice. Or if your data has consistency and requires shigh-security data protection, you can choose Oracle or other RDBMS.

REFERENCE

Arsenault, C. (2019). *The Pros and Cons of 8 Popular Databases - KeyCDN*. [online] KeyCDN. Available at: https://www.keycdn.com/blog/popular-databases [Accessed 20 Oct. 2019].

Arteaga, A. (2019). *Oracle Definition, Functions, Advantages and Disadvantages | Symantec Connect*. [online] Symantec.com. Available at: https://www.symantec.com/connect/articles/oracle-definition-functions-advantages-and-disadvantages [Accessed 20 Oct. 2019].

Claburn, T. (2019). *How to secure MongoDB – because it isn't by default and thousands of DBs are being hacked*. [online] Theregister.co.uk. Available at: https://www.theregister.co.uk/2017/01/11/mongodb_ransomware_followup/ [Accessed 20 Oct. 2019].

DataFlair Team (2019). *MongoDB Data Modeling with Document Structure - DataFlair*. [online] DataFlair. Available at: https://data-flair.training/blogs/mongodb-data-modeling/ [Accessed 20 Oct. 2019].

Docs.bmc.com. (2019). *Hardware requirements to support small, medium, and large environments - BMC TrueSight Infrastructure Management 10.0 - BMC Documentation*. [online] Available at: https://docs.bmc.com/docs/display/public/tsim10/Hardware+requirements+to+support+small%2C+medium%2C+and+large+environments [Accessed 20 Oct. 2019].

Hackernoon.com. (2019). *MongoDB vs MySQL Comparison: Which Database is Better?*. [online] Available at: https://hackernoon.com/mongodb-vs-mysql-comparison-which-database-is-better-e714b699c38b [Accessed 20 Oct. 2019].

Itconsultantsinbayarea.com. (2019). *Some Pros & Cons of Selecting Oracle as your Database Platform | Diamond Technology*. [online] Available at: http://www.itconsultantsinbayarea.com/some-pros-cons-of-selecting-oracle-as-your-database-platform/ [Accessed 20 Oct. 2019].

Kirkpatrick, D. (2019). *Mongodb - Security Weaknesses in a typical NoSQL database*. [online] Trustwave. Available at: https://www.trustwave.com/en-us/resources/blogs/spiderlabs-blog/mongodb-security-weaknesses-in-a-typical-nosql-database/ [Accessed 20 Oct. 2019].

Noworyta, T. (2019). *The pros and cons of using MongoDB*. [online] Medium. Available at: https://medium.com/selleo/the-pros-and-cons-of-using-mongodb-8e7601857a8e [Accessed 20 Oct. 2019].