

## 1.

1.1 What is driving the increasing use of data warehousing? It is now the norm for data warehouses to store terabytes of data.

Every business needs to increase their profits and make better decisions. With the evolution of using social media and electronic mediums for advertising, marketing and receiving customer feedbacks, businesses generate humongous amount of data every day. These data when cleaned and summarized generates valuable insights in to market and customer behavior as well as weaknesses and strengths of company policies. Datawarehouse is the solution to store and process such big data to generate reports and information about the businesses. With the advancements of data storage and manipulation technologies like cloud platforms, parallel processing and data mining more and more organizations and businesses choose to have a data warehouse of their own as a valuable source of business intelligence.

1.2 How is the data warehouse used in Business Intelligent (BI) solutions?

A production database focuses on production tasks and daily processing of transactions. A data warehouse contains current and historical data stored according to a set of constraints and rules and is able to process complex queries on big amount of data to generate information. The data model of a data warehouse is designed to analyze data so their tables and joints are denormalized and less complex (Tobin, 2020). Whereas production database has complex joins and tables. Data warehouses handle few users simultaneously while production databases can handle thousands. A data warehouse uses Star, Snowflake and Galaxy schemas while production databases use relational database schemas. Since data warehouses are for analyzing big data and generate information, they process complex queries with longer time while production databases handle large amounts simple queries quickly. Granularity in databases means the level of detail or summarization of the units of data in the data warehouse (Dhakal, 2014). Production databases are usually low granular and data warehouses keep data at a more granular level as they need to generate reports and information out of them.

For an example:

- A. Order\_lines(order\_num, item\_num, quantity, cost)
- B. Order\_lines(order\_num, item\_num, cost)

Representation A is used by sales and B is used by marketing. Thus “cost” in B is actually quantity\*cost. Hence the same attribute has different granularity in different tables.

Data warehouses are made with the intention to process large amounts of data to uncover useful patterns and information about businesses hence data warehouses provide business intelligence to businesses through aiding the persistence of big data that is to be mined using mathematical methods.

### 1.3 What is meant by the term “data mining”?

Data mining converts large batches of raw data into useful information. It is literally the process of mining large amounts of raw data to excavate useful gems of patterns, correlations, information and finally acquiring knowledge out of them.

## 2.

### 2.1 Measures that could be used to manage user account passwords.

Using two-factor authentication can secure the passwords as it is secured with a second layer of authentication. Using password managers is another good measure. Awareness is key in managing passwords as well. Sharing, writing down passwords should not be done. Using strong combination of characters for passwords makes attackers' job hard. When storing passwords, we must hash them; they should not be stored as plain text. Using cryptography for storing passwords is a very good method to manage passwords.

### 2.2 What would happen when the trigger is fired? Discuss how this can affect normal database operations.

The trigger inserts a new row to the table DEPT\$audit which holds auditing records, each time a new record is inserted, deleted or updated in the DEPT table. The change\_type column holds the type of change; Update, Delete or Insert. The attributes, changed\_by and changed\_time hold who did the change and when. If the change is an Insert, the attributes DEPTNO, DNAME, LOC of audit\_data row will be populated with the new values. If the change is a delete or a update, then the attributes will hold old values of the corresponding values of attributes from the DEPT table.

How the triggers will affect the normal operations depends on many factors. How often the DEPT table is used, how many triggers will be issued within certain period of time, the memory available in the system, how much the other process use the memory and also how much of space will it take to store all the audit logs compared to vacant memory for the other persistence processes and also the execution plan of the trigger will determine how the trigger will affect the normal operations.

### 2.3 Top five steps to a secure Oracle Database Server.

1. Locking down the default privileged accounts.
2. Downloading Security Patches and Contacting Oracle Regarding Vulnerabilities.
3. Discourage users from using the NOLOGGING clause in SQL statements.
4. Require all database connections to use a strong SID.

5. Enable data dictionary protection.

## References

Dhakal, A. (2014) *Modelling a Data Warehouse - Data Science Central*. Available at: <https://www.datasciencecentral.com/profiles/blogs/modelling-a-data-warehouse> (Accessed: 29 July 2021).

Tobin, D. (2020) *Data Warehouse vs. Database: 7 Key Differences*, Xplenty. Available at: <https://www.xplenty.com/blog/data-warehouse-vs-database-what-are-the-key-differences/> (Accessed: 29 July 2021).