Database administrator (dba) research papers example

Business, Management

Database

Data is a term that represents the raw facts about an entity about which a person or organization may have need to store. A computer user feeds this data into a computer using input devices such as a keyboard or scanner, where the data gets manipulated by specialized computer programs. This is called processing after which meaningful and well organized data results. This is called information. With time an organization accumulates large amount of data and information.

The USG (2014) asserts that an organization must therefore store this data and information into an organized and efficient system where users may access it easily and fast, and use it when need arises. This is what constitutes a database. When an organization performs this process using computers and related software, the result is an electronic database.

A database can be as simple as a book catalogue in the local library, maintaining a simple inventory of the books in the library and the keeping the borrowing and return records, or a simple filing system in an office. This is usually a manual database that can be managed easily by an employee in the organization. However, in many large modern organizations, the use of information and communication technology (ICT) demands that the system maintains an electronic database. An electronic database works within a specific model which defines its logical structure and depicts the relationships that exist between the data in the database. KUAS (2014) identifies three distinct types of electronic database models namely Network, Hierarchical, and Relational. The hierarchical model breaks down data progressively into a tree structure. A network model interlinks different

elements of data in such a way that data items become accessible through others.

The most commonly used model is the relational model which represents data in form of two-dimensional (2D) tables called relations. According to KUAS (2014), the data table consists of rows and columns. Rows represent one complete set of characteristics that form that data item or entity. Rows are also called turples or records or entities. Columns represent each individual characteristic of an entity. They are also called fields or attributes. The programs or software that manage the database in terms of data input, validation, storage, extraction, reporting, and so on are called database management systems or software (DBMS). In a relational environment, they adopt the name relational database management systems (RDBMS). For the purposes of this paper, we will place a bias on the RDBMS.

Modern databases maintain huge amounts of data, for example, a bank that has branches spread across a country, a continent and even all over the world, coupled with a global automated teller machine (ATM) network would have millions of customers performing millions of transactions daily. This would result to a huge and complex database that maintains the customers' details, all ATM transactions details, counter transactions, loans records, interests and fines, bank administrative records such as accountancy, payroll, investments, and so on. A database of this magnitude and complexity requires the attention of a database professional called a database administrator (DBA).

Structured Query Language (SQL)

KUAS (2014) describes SQL as a computer programming language that serves as a standard across a diversity of databases in the manipulation of the data within a relational database. SQL is mostly inbuilt into database management software (DBMS) such as dBASE, Microsoft Access, and so on, although there are independent versions of it such as MySQL. The language has received global acclamation for example from the American National Standards Institute (ANSI) and International Standardization Organization (ISO) because of its seamless cross-system ability to manipulate data inside relational databases.

According to KUAS (2014), SQL operates using six specific paradigms. These include;

- DELETE – this operation removes the record that meets a given condition from the table.

- UPDATE – this modifies the value for a record that meets a given condition with the new (provided) value.

- PROJECT – this creates a new relation with values extracted from another relation.

- JOIN – it creates a new relation by combining data from two relations that share common characteristics.

- UNION – this provides the ability to extract data simultaneously from two relations that share common attributes.

- INTERSECTION – this operation creates a new table from two different tables that share common attributes, in such a way that the two tables become accessible through the new intersection relation.

- DIFFERENCE – the operation compares two tables with the same attributes exactly. It then extracts all the records from the first relation that do not exist in the second relation, and stores them into a new table.

A DBA according to Timothy O'Leary and Linda O'Leary (26) is a well-trained and experienced computer scientist who holds a minimum of a university degree in computer science. In a nutshell, the responsibilities of a DBA include identifying the most effective and efficient system for the storage, organization, and easy and fast access to the organization's database. Mark Spenik and Orryn Sledge (2001) emphasizes that the DBA ensures that the database server is available always, and that a DBA would most probably have a database design and administration background.

The DBA has a major responsibility to protect the database. According to Robidoux (2014), the most effective method is the Disaster Recovery Plan (DRP). The DBA evaluates the threat to the database progressively downwards i. e. from the server level, to the database level, and finally to the table (relation) level. Again, since relational databases constitute several relations within the database, and such databases may be spread out across several servers, the DBA must prepare a priority list showing which server or database would be affected first and the order of events trickling downwards. The next step is to identify the plan of action to take on each of these levels in case the disaster struck. After identifying the threats and the plan of action, Robidoux (2014) recommends carrying out of tests to ensure that the plan actually works.

The DBA also institutes security mechanisms for the protection of the database against un-authorized access, data corruption, data integrity and

loss. He or she is also responsible for making copies of the database at reasonable intervals, a process called data backup, so that the system remains secure even in the event of mishap such as data corruption, hacker attack, virus infection, and so on. Robidoux (2014) cautions that an effective backup system should consist of daily full backups and transaction log back up every 15 minutes or every one hour, or a full daily backup depending on the sensitivity and complexity of the system and its data.

Backups constitute security against data loss. Microsoft (2014) asserts that the DBA can create backups through bulk copying. He or she can perform this using the SQL command-line tool BCP or BULK INSERT. The DBA may also delegate this work to a trusted user by allocating them a server-login account, and giving them read and write (R/W) privileges. Again, Microsoft (2014) suggests that the DBA may also have the responsibility of applying SQL hot-fixes. In the event that he or she needs to delegate this work, the DBA must provide the user with an account in the Local Administrator group.

The DBA must ensure the security of the system and manage the same through the security levels. Robidoux (2014) asserts that these levels are available both at the operating system (OS) level and the SQL server level. To fully protect the database from unauthorized access such as doorways, the DBA must implement access permissions at both level. Operating systems allow at least three access levels which include administrator (full access), standard users (limited access to OS features but full access to applications) and guest (limited access to user applications only). According to Robidoux (2014), the DBA should also implement SQL Server roles which

set out permissions for SQL Server access, database access, and table access. The DBA must ensure that users have only as much access as they need.

Microsoft (2014) provides a unique angle to the issue of user permissions by stating that the DBA must institute permissions for the startup account for SQL Server and the SQL server agent. Since the SQL Server and the SQL Server agent startup accounts can belong to the Domain Users group, the DBA may exempt them from the powerful local administrator group. This way the DBA makes both effective users of the database but devoid of the power of the database administrators. The DBA is also responsible for starting up and shutting down both the SQL Server and SQL Server Agent Services.

The DBA also has the responsibility of running system and data integrity checks daily. According to Robidoux (2014), SQL servers are constantly monitored by DBAs and offer utilities for performing integrity checks such as DBCC. Microsoft (2014) informs DBCC is the process checking the consistency of databases. The tools are effective in checking logical integrity of data including allocation and structure. They produce reports that the DBA can study in order to identify any issues that demand the DBA's attention. Examples of those tools include DBCC CHECKTABLE, DBCC CHECKDB, and so on.

The other sensitive responsibility of the DBA is maintenance of indices (indexes) created in SQL databases. An index is a database tool that speeds up retrieval of data by pointing at specific data pages in the database. This way, the search is constrained to the pages pointed at rather than an

attempt at searching the entire database. New indices are fast and efficient but with time they become fragmented and slow down the search process. Robidoux (2014) asserts that the DBA can perform index maintenance through maintenance plans that perform routine rebuilding of indices and defragmentation of indices. Index defragmentation is the process of identifying all the fragments of each index, combining them to form a single index, rewriting the new index into a fresh location and deleting all the fragments.

SQL-based systems such as the SQL server constantly write event logs such as error logs every time the error occurs. It is the responsibility of the DBA to monitor and review these logs in order to identify areas that need attention and correction. Robidoux (2014) identifies the SQL server error log as the most common log that DBAs target. The other logs that a DBA must review include the operating system log, SQL server database mail, and SQL server agent. Reviewing these logs help the DBA to trace the source of errors and identify the nature of that error, and this way he or she will manage to correct the error. Logs identify events such as time of the event, backup information, database integrity, disk integrity, and so on.

The other responsibility of the DBA is the management of the SQL Server Agent job scheduler. Robidoux (2014) asserts that this agent is useful in the automation of several DBA responsibilities. If the DBA manages the SQL Server Agent well, then it may take away a huge load of his or her shoulders. The agent can automate data backup, integrity checks, rebuilding of the index, and so on. The DBA must allocate the jobs for automation to the agent and constantly monitor that the tasks are performed successfully every time.

Another important role of the DBA is monitoring of the performance of the system, and not just the database. Sometimes the RDBMS may not be operating at their best, and this creates the impression that the database is not efficient. The DBA must monitor the applications, which comprise mainly of SQL operations, to ensure that they are operating optimally. Robidoux (2014) suggests that the DBA has useful SQL at his or her disposal which includes Database Engine-tuning Advisor, Index Tuning Wizard, performance monitor (Microsoft 2014), and so on.

According to IBM (2012), the task of monitoring performance includes analyzing SQL statements in order to determine where there may be delays in execution, and identifying when the database performs rollback on large transactions thus adversely affecting other transactions. Again, sometimes the system experiences a lock. The DBA identifies the source of locks by performing traces on SQL statements, and identifying slow SQL executions.

Works Cited

O'Leary, Timothy, J. and Linda I. O'Leary. " Databases". Computing Essentials, 2008. Web.

5 May 2014.

Microsoft, Microsoft SQL Server, 2014. Web. 7 May 2014.

National Kaohsiung University of Applied Sciences (KUAS), Databases, Foundations of

Computer Science Cengage Learning, 2014, Web. 3 May 2014.

Robidoux. Greg. SQL Server DBA Database Management Checklist. 2014.

Web. 7 May 2014.

Spenik, Mark, Sledge, Oryn. " What is a Database Administrator", Sams

Publishing, 20 March

2001. Web. 7 May 2014.

2014. Web. 3 May 2014.