

The Database Administrator

All databases need an administrator. The database administrator (usually known as the **DBA**) is responsible for the database system itself and the design, definition and control of all the data it stores. The DBA is not necessarily a single person; for a small, single user database that is used only within working hours the DBA role could be filled by someone with other duties who spends the occasional afternoon or evening on the task whereas in a large database which runs day and night and with many users in many different locations the DBA could be the head of a large team. The various tasks of the DBA are discussed in more detail below.

Access Control

Firstly the DBA must ensure that only authorized users access and change the data and then only the data they need to access and change. This means that the DBA has to provide passwords where necessary, appropriate views for each user and probably input screens for data entry.

Session Control

The work patterns within the organisation must be arranged so that the database system gives priority to the right things at the right time. For example some programs can run just as well in the middle of the night as in the daytime whereas during working hours everybody wants the database to run as though they were the only users. This cannot normally be arranged by simply setting up the database system in the right way; it usually involves dealing with people instead.

Training and Consistency Controls

New users have to be trained to use the database system and so do new members of the database administrators' team. The correctness of database system data is usually ensured by using programs to access this data and these programs check that the data is not corrupted.

The database can only be kept consistent if the programs that access it work in a way that ensures that, provided the database is in a consistent state when they start to run, it will be in a consistent state when they finish. It is the DBA's responsibility to provide an environment that ensures that new or modified application programs are correct. And, if anything goes wrong, it is the DBA's responsibility to find the errors and repair them.

Database Design

The initial database definition is generally relatively easy; problems occur when things change. Monitoring change is difficult but any sort of redefinition or evolution of the system has to be carefully documented and the system has to be watched for unexpected side effects. Of course the database administrator is also responsible for ensuring that there are no unexpected side effects of changes, but everybody makes mistakes.

Changes sometimes occur because the way the database is used alters but more often it is the real world that changes and, as the database system reflects the real world, it has to be changed too. Large scale database systems have built in tools to help the DBA and provide some sort of automatic record of changes.

Recovery Control

Things go wrong for all sorts of reasons. We must be able to recover to a consistent state after a failure and be able to move forward to where we should have been. In a large, multi-user database most of this is managed by a program called the **recovery manager** but the recovery manager will only work if the database's back up systems can be relied on and it is the DBA's responsibility to see that this happens. In a small database there may be no automatic back up available and here the role of the DBA is even more important. If data is valuable enough to store in a database then it ought to be protected from disasters.

Storage Media Control

This concerns the allocation of database tables to existing storage devices; the relationship between the logical representation of the database and the physical devices it is stored on. The DBA is responsible for a choice of storage strategy that optimizes the performance of critical applications without degrading the performance of other applications.

The DBA is also responsible for the indices used by the database system. An index allows the user to find a particular row or group of rows quickly and to produce ordered lists of data efficiently. They do have disadvantages; indices speed up access to database tables at the expense of slowing down any changes to the data and they take up a lot of physical storage space.

Finally the DBA has to deal with changes in the physical storage available. These can occur either because the data has outgrown its physical storage space or because a new storage device or medium becomes available. Having more space might not normally be considered a problem but when there is no possibility of unloading and reloading the database it is.

Monitoring and Statistics

The DBA must keep track of the population of tables and discs; that is the number of rows stored in them. He or she must be aware of the distribution of data and look out for skewness and patterns. If 5% of the discs (or whatever storage is used) are almost full and the rest are almost empty then the DBA has probably made a mistake. Similarly the DBA must keep track of the usage of discs and DML programs. If everybody wants data on the same disc at the same time, while all the other discs are unused, then the DBA has got it wrong again.

Documentation

Everything has to be appropriately documented and the documentation has to be maintained. This is particularly important if the whole DBA team are in the same lottery syndicate.