

Relational Database Design

Dinesh

Database Management System (DBMS)

Database is collection of interrelated data and Database Management System (DBMS) is collection of database and programs to facilitate storing, accessing and manipulating data, creating new schemas, modifying /removing existing schemas.

Database Management Systems are developed to manage large amount of data existing in the organizations. It can be used by individual also for storing data about his/her own activities like daily expenses, daily interest calculations. To manage large amount of data, DBMS facilitates to define structure of storages and mechanism to manipulate the data.

In addition to storing data, DBMS also provides mechanism to secure stored data, mechanism to make backups of the existing data such that data can be recovered in case of data lost or some kind of problem.

Database stores database about entities, their attributes and relationship between entities. Entities are stored in tabular form in database in relational database. Database provides very fast mechanism to access data, it uses concept of indices to fast retrieval of data. Indices are maintained in database itself.

Database has data dictionary which stores information about the database itself. Data dictionary store information about the created relations, their relationships, constraints. Whenever changes have to be made in database tables, data dictionary is consulted. At the time of inserting new data also to check consistency of provided data, data dictionary is used.

Consistent database is that one in which all the data obey the constraints specified for the database. There can be different constraints on database tables, attributes of tables (relations). Let' consider employee table with attributes employeedid, name, age, post. Each of these attributes must have data type, that is what type of value each attribute can take, (values can be string, number, real number etc. also string with limited number of characters, numbers within range etc). And each attribute may have constraints (rules) also. Here, for employee, value of age can't be 300 as nobody lives up to 300 years, similarly, value of post can't be that value which does not exist in the organization. And table which have only allowed values is consistent.

Similarly, in case of transactional database, consistency of database may be violated by updates, even the values in the columns of table obey specified domain. For example, let's consider bank database. Bank database contains account tables. Let A and B be two account with 1000 and 1000 each. The total amount of A and B is 2000. If we transfer 500 from A to B, then then sum of A and B still must be (500 + 1500) 2000. But if somehow the deduction from A happened but addition to B did not happen, in this cased A will have 500 and B will have 1000 and total will be 1500 which is wrong. This means the performed transaction made database inconsistent. Here, both values obey data types of columns, but the sum is incorrect after transaction.

[Transaction is set of operations which work on database, it consists of one or more SQL statements. SQL is structured query language used to work with database]

Applications of Database System

Relational Database Design

Dinesh

Data is one of the most important component for every organization and individuals, information can't be generated without data. So, it has become top most priority for all the organizations to store data properly such that data/information can be retrieved whenever required.

Organizations like banks, airlines, universities, hotel, hospitals etc all have some kind of data requirements and they also produce data from their daily activities. These organizations need to store those data properly for information generation, future references. Here application of database for organizations in storing data about the internal activities of the organizations.

Banks have several different types of accounts and clients for those accounts; those accounts have their own features. Banks use database to store data about their clients, accounts etc. The database of the bank enables bank to know number of clients, amount deposited each day, amount withdrawn each day etc.

Is there any other mechanism to store data?

Obviously, there is another mechanism to store data and that is using file system. Before the development of DBMS, organizations used to store data in the files. But there are lots of limitations with the file system. The major limitations with the file system to store data are as follows:

- Data redundancy and inconsistency
- Difficult in accessing data
- Data isolation
- Integrity problem
- Atomicity problems
- Concurrent access anomalies
- Security problems

These limitations of the file system prompted the development of database management systems. The DBMS has no limitations as seen in the file system as DBMS has included all the features to cope with those mentioned limitations of the file systems.

Three level architecture of DBMS

DBMS is developed to remove the entire problem that users have to face with the file system for storing data. So, to use system without problems of data redundancy, data inconsistency, and ease of sharing, ease of security and backup, ease for concurrent access, the complexity of the DBMS has been hidden from the users and only the abstract view of the database is provided to the users. To hide underlying complexity and gain ease to use interface for the user, three level architecture has been developed. This architecture hides certain details like how data are stored and maintained in the storage devices.

The three levels are:

- **Physical level:** here how data is stored is defined
- **Logical level:** here what are the data and how data are related with each are defined
- **View Level:** this level provides view of data stored in the DBMS to the users, one database can have different views and as many views as users requirements.

Instances and Schemas

Relational Database Design

Dinesh

Data is stored in the database, with the time volume of data, values of data will change. But at specific moment there will be only one view of the database and that is known as the instance of the database, with the time instance of database will change.

Database Schema is the overall design of the database and it does not change frequently like instance of the database. Physical level has physical schema to describe the database design of that level, and logical level has logical schema to describe the database design of that level. View level may have several schemas.

The logical schema is the most important one as programmers construct applications using the logical schema. The physical schema is hidden under the logical schema. Application programs can be developed without affecting or making changes to the physical schema, this ability of application programs is known as physical data independence.

Data Models

Data model is collection of tools used to describe data, entities, relationship between entities, semantics of data, rules for making data consistent. The structure of the database depends on the data model. Two most popular data models are: Entity relationship model and Relational model.

Entity Relationship Model (ER Model)

Entity relationship model shows the pictorial representation of database. That is it shows the entities about which data have to be stored, relationship between the entities, constraints on the entities and relationships between entities. ER diagram uses symbols to show entities, attributes, relationship between entities, connections etc.

- **Rectangle** is used to represent entity
- **Diamond** is used to represent relationship
- **Oval** is used to represent attribute
- **Straight line** is used to connect attributes and entity, entity and relationship

Entity is a real-world thing/object (abstract or physical) which have some properties (characteristics) and which is related with one or more other entities.

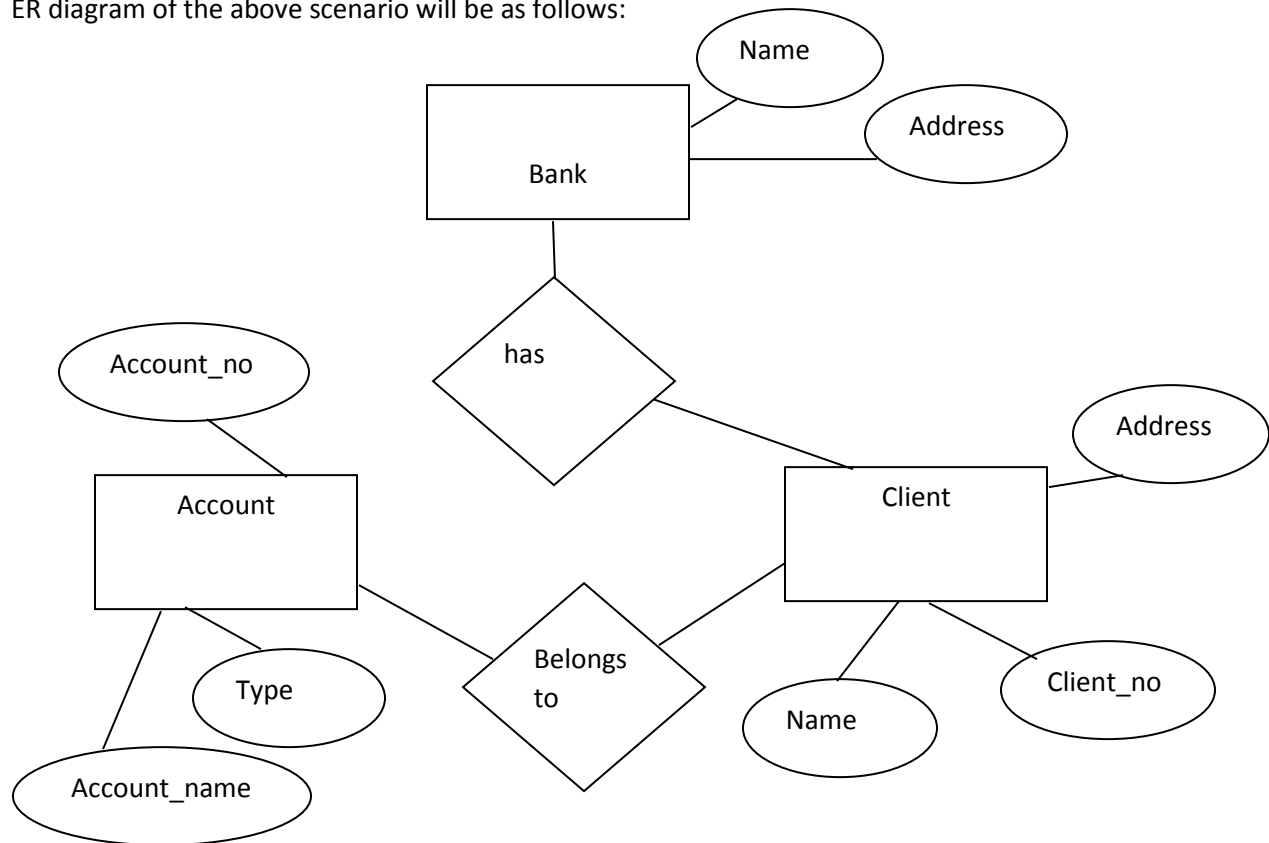
Let's take a simple scenario of bank. Bank is an organization where money of clients is deposited, bank also provide money as loan to the clients. Clients create accounts in bank to store money or borrow loan. Account is described by properties like: account_no, account_name, type, client is described by client_no, name, address.

Here in this simple scenario of bank we can find entities Bank, client, account etc. Bank needs to store data about these entities. These entities are somehow associated with each other. Here client is associated with account, bank is associated with account and client both.

Relational Database Design

Dinesh

ER diagram of the above scenario will be as follows:



Relational Model

Relational model is the most popular model for creating database. It stores data in the table, known as relations, each table (relation) represents one entity (generally), columns of table represent attributes of entity, rows of table represent members of the entity.

In the relational model constraints for the entities can be easily specified.

Id	Name	Address	ContactNo
C101	Geeta	New Baneshwor	4444444
C102	Ishwor	Old Baneshwor	4444441
C103	Jeevan	Sano Thimi	4444442
C104	Usha	Balaju	4444443

Here is a simple table for storing data about the persons; persons' id, name, address, contactno is stored in the table. Each person is uniquely identified by the attribute Id. In the given table there are four rows, these four rows represent four different persons with having value for attributes: Id, Name, Address, ContactNo. Each row of table is also known as **tuple**.

Relational Database Design

Dinesh

Database Language

Database facilitates to store, access and manipulate data. Database tables are created to store data. Structured Query Language (SQL) is used in DBMS to create schema of tables and manipulate schema of table, insert new data, manipulate data. There are two parts in SQL:

- Data Manipulation Language (DML)
- Data Definition Language (DDL)

Data Definition Language (DDL):

DDL is used to create, alter and drop the schema of relations. SQL has several data types that are used to specified type of values attribute can take. Some of the data types are varchar, int, char etc. Data type specifies the set of possible values for attribute. The statement of SQL used to create, modify and remove database relations are:

- **Create:** This statement is used to create relations.
- **Alter:** This statement is used to modify existing relations. New attributes, constraints can be added and attributes, constraints can be removed from existing relations
- **Drop:** This statement is used to remove the existing relations.

Data Manipulation Language (DML):

DML is used to insert, retrieve, remove, manipulate existing records (rows) of tables. The statement to enter record, retrieve record, remove record and manipulate record are as follows:

- **Insert:** This statement is used to insert new record to the relation. The name of attributes and values for the specified attributes are specified in this statement.
- **Select:** This statement is used to retrieve records from the relation. Condition can be specified while retrieving records. In select statement where clause is used to specify condition/criteria.
- **Delete:** This statement is used to remove records from the relation. Condition can be specified while retrieving records. In delete statement where clause is used to specify condition/criteria.
- **Update:** This statement is used to modify existing records of relations. Set keyword is used in the update statement and where clause is used to specify condition/criteria.

Database storage and Querying

Several modules are developed in database management system to ease storing of data and retrieving data. The major two functional components of database system are: storage manager and query processor.

Storage Manager:

Storage manager facilitates the application program by providing interface to interact with the data stored in the lower level of the database. The storage manager interacts with the file manager. The raw data is stored in the disk by the file system provided by the operating system. The storage manager does

Relational Database Design

Dinesh

the function of storing, manipulating and retrieving data. The storage manager implements different data structures like data files, data dictionary, indices. The storage manager has following components:

- **Authority and integrity manager:**
 - Integrity manager component is responsible for maintaining consistency of database. It helps to impose integrity constraints on relations (tables) and maintain imposed constraints. Authority manager helps to create users, define role and set of actions users can take on the database tables. It helps to make database secure by defining roles and functions for the users.
- **Transaction manager**
 - This component is responsible for executing transactions (set of operations on database) and maintain consistency of database after the execution of transactions. Transaction takes consistent database, works on it and returns consistent database. Transaction manager maintains consistency of database even database is accessed by more than one transaction. There will be concurrency control mechanism to check concurrent transactions and maintain consistency of database.
- **File manager:**
 - File manager manages the disk space required to store database and data structures required to represent data stored on the disk.
- **Buffer manager**
 - Buffer manager is responsible for transferring data between disk storage and main memory and decide which data to keep in main memory and which not to. Size of database can be very large than main memory, here buffer manager plays big role in handling database larger than size of main memory.

Storage manager uses different data structures to manage different functions:

- Data files: is required to store database itself
- Data dictionary: stores data about database
- Indices: required for fast access of data.

Query processor

Query processor simplifies the data access process. It keeps away users from the burden of understanding how the data is stored in the disks. It translates queries written in high level language at the logical level. Query processor has following components:

- DDL interpreter
 - This component interprets DDL statements and stores the DDL actions in the data dictionary.
- DML compiler
 - Translates DML statement into evaluation plan. DML compiler creates more than one evaluation plan of a query in terms of time, memory and also picks the best plan in terms of time and memory for execution.
- Query evaluation engine
 - Executes the best plan to execute query generated by DML compiler.

Relational Database Design

Dinesh

Database Architecture:

The architecture of database is influenced by computer systems on which database system runs. Database system can be centralized or client server. The architecture of database depends on how the users will communicate with the database server, users directly communicate with the database server or there is some application program between users and database server.

Two-tier architecture:

In case of two-tier architecture application programs resides within the client machine. Some of the popular application program interface are ODBC and JDBC which are used to interact between client and database server.

Three-tier architecture:

In three-tier architecture, client machine only has user interface and does not interact with database directly. Here client works only as front end device. Client communicates with application server through the front end (User Interface), and the application server communicates with the database server to access data. In the application server business logic is embedded which specifies what actions to carry out in which condition.

For larger applications three-tier architecture is more suitable as changes if required to be made, should be made only to the application server. But in case of two-tier architecture if change has to be made in application programs, then change must be made in all the client machine as each client has its own application program.

Database user

Database is developed for storing new data and retrieving existing data. For the purpose of storing, manipulating, retrieving and developing database several types of users will be in work who can be categorized as database users or database administrators.

Database users:

- **Naïve users:** the users who will be interacting with database using existing application programs are Naïve users. For example, users checking their email online are naïve users.
- **Application Programmers:** Application programmers are the users who write programs like user interface to interact with the database. These users facilitate naïve users to interact with the database and access data and reports.
- **Sophisticated users:** These users directly submit database query written in query language to interact with the database. They may also use data analysis tools/software to explore data.
- **Specialized users:** These users are sophisticated users who write specialized database programs like expert systems, knowledge based systems etc.

Database Administrator (DBA)

Database Administrator is person who has central control over the whole database. He or she can make changes, add new users, add new relations/tables, constraints in the database, alter schema of database. DBA has responsibility of maintaining database such that the performance of database is not degraded, also should make provision of backing up database and recovery mechanisms.

Relational Database Design

Dinesh

The functions of DBA are:

- Schema Definition
 - Create schema of database by executing DDL statements
- Storage Structure and access-method definition
 - How data will be stored and how data will be accessed?
- Schema and physical-organization modification.
 - Altering existing schema of database whenever required.
- Granting of authorization for data access.
 - Defining users and their roles, specifying and grants roles and rights.
- Routine maintenance.
 - Backing up database, tuning database, removing unnecessary relations, views.

Questions:

- What necessitated database?
- Where database can be used?
- How database management system eases the storing and retrieval of data?
- Explain different types of database users.
- What is query language?
- Explain the components of storage manager.
- Explain the components of query processor.
- Differentiate between two tier and three tire architecture.
- What is the importance of Entity relationship diagram?
- What are the advantages of relational model?