UNIT 01: Introduction to database

Lesson 1.2: Database System Concepts and Architecture

STRUCTURE:

1.2.1. Database System Concepts And Architecture

1.2.2. Three Schema Architecture And Data Independence

1.2.3. Database Languages And Interfaces

- DBMS Languages:
- DBMS Interfaces:

1.2.4. Database System Environment

1.2.5. Centralized And Client /Server Architecture For DBMS

1.2.6. Classification Of Database Management Systems

1.2.1. Database system concepts and architecture

In Client /Server architecture, the functionality of the system is distributed into two modules. Normally application programs reside in **Client module .**The client module handles user interaction and also provides graphical and form based interfaces. **Server module** handles storage of data, search, access and other functions.

DATAMODELS, SCHEMAS AND INSTANCES

Data model is the collection of concepts that can be used to describe the structure of database. It facilitates to achieve abstraction. Structure of the database means the types of data, constraints and relationships that should be applicable to the data. Data model also includes set of basic operations for stating retrievals and updates.

Categories of data models

Based on the types of concepts used to describe the database structure the data models can be categorized in to

High –level or conceptual data model:

It provides concepts that are close to the way many users observe data. It uses ideas such as entities, relationships, and attributes. Entity is the real world object such as employee or project in database. Attribute represents some property of interest that describes an entity such as name or salary. Relationship between two or more entities represents an association among the entities.

Low –level or physical data model:

It provides the concepts that describe the details of how data is stored in the computer.

Representational or implementation data models:

It provides concepts that can be understood by end users but that are not too far removed from the way data is organized within computer. These include relational data model, network and hierarchical data models. Representational data models represent data by using record structures and so called record based data modes.

Schemas instances and database State

The description of the database is called database schema. It will be specified at the time of database design and will not change repeatedly. The displayed schema is called schema diagram. We call each object is called a schema construct. Schema Diagram for university database is as follows

STUDENT

Name	StudentNumber	Class	Major

COURSE

CourseName	CourseNumber	CreditHours	Department

PREREQUISITE

CourseNumber	PrerequisiteNumber

SECTION

SectionIdentifier	CourseNumber	Semester	Year	Instructor

GRADE_REPORT

StudentNumber	SectionIdntifier	Grade

Figue1.4

A Schema Diagram gives only some aspects of the schema such as name of the data items and some of the constraints. Other aspects such as data types and various relationships among those data items are not there in the schema diagram. The data in a database at a particular moment is called database State. It is also called current set of occurrences in the database. Many database states constructed to correspond to a particular schema. Each schema construct has its own current set of instances in a database state. The DBMS stores the description of the schema constructs and constraints (metadata) in the DBMS catalog.

1.2.2. Three schema architecture and data independence

In this architecture schemas are defined in the following three levels

1) Internal level:

It has an internal schema and describes the physical storage of the database. It uses physical data model and gives the complete storage details.

2) Conceptual level:

It has a conceptual schema and describes the structure of the whole database. The conceptual schema hides some details of the physical storage and describes entity, data types, relationships, user operations and constraints. Representational data model is used to describe the conceptual schema

3) External level:

It includes external schemas or user views. External schema describes part of database that a particular user group is interested and hides rest of the database from users.

Each user group has its own external schema. DBMS transforms the requests specified in the external schema in to the request corresponds to the conceptual schema in to the request in to the internal schema. In case of retrieval operation the data fetched from the database is reformatted to match the user's external view. The procedure of transforming requests and results between two levels is called **mappings.** The following figure shows the pictorial representation of three schema architecture.

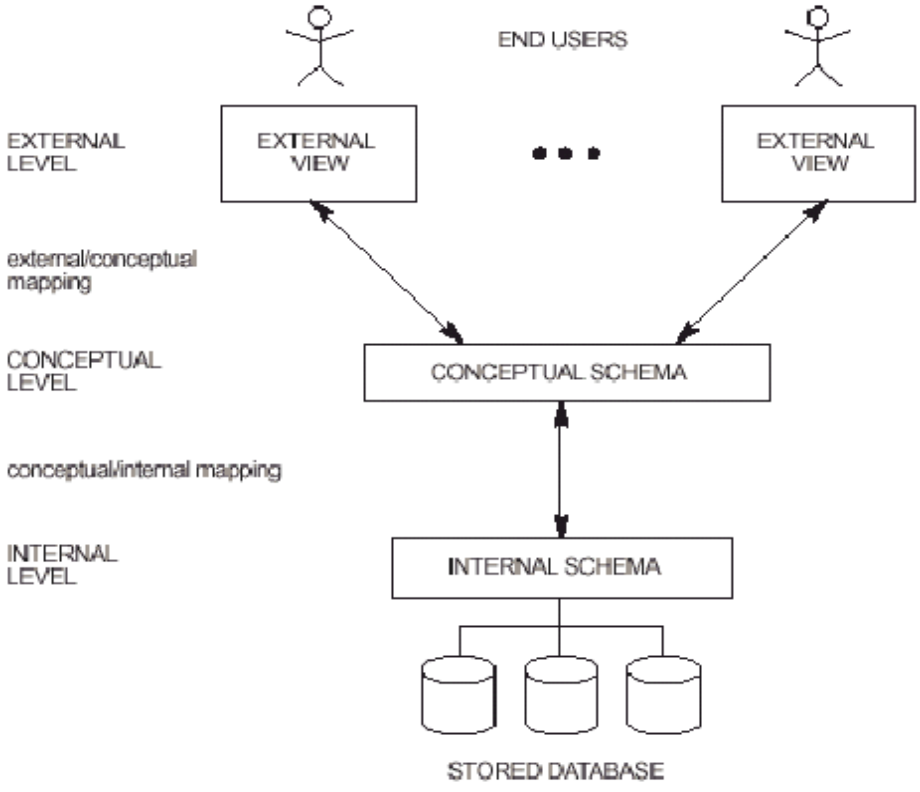


Figure 1.5

Data independence

The three schema architecture is used to describe the concept of data independence. There are mainly two types of data independence**.**

Logical data independence:

It is the capacity to change the conceptual schema without changing the external schemas or application programs. We can change the conceptual schema by adding new record type or data item, to change the constraints or to reduce the database by removing a record type. Only view definition and mappings need to be changed in DBMS to support logical data independence.

Physical data independence:

It is the capacity to change the internal schema without changing the conceptual schema and external schema. Changing the internal schema may be needed after reorganizing some physical files.

1.2.3. Database languages and interfaces

DBMS provides appropriate languages and interfaces for each user groups.

- **DBMS Languages:**

In many DBMS s when there is no separation between levels **Data Definition Language (DDL)** is used by database designers to define schemas. When there is a clear separation between conceptual and internal levels DDL is used to define Conceptual Schema only. **Storage Definition Language (SDL)** is used to define internal schema. The mappings between these languages can be specified in any of these languages. View Definition Languages (VDL) is used to spiffy user views and their mappings to conceptual schema. DBMS provides **Data Manipulation Language (DML)** for specifying manipulation operations such as retrieval, insertion, deletion and updation of data. There are two types of DMLS.A **High-level or nonprocedural DML** which is used on its own to specify database operations in a concise manner. A **low level or procedural DML** which is embedded in general purpose programming language. Low level DML is also called record at a time DML retrieves each record and processes separately.

- **DBMS Interfaces:**

Menu Based Interfaces For web Clients or Browsing:

The interfaces provide menus that helps user through formulation of the request. Pull down menus are a well-known technique in web based user interfaces.

Form Based Interfaces:

It displays a form to each user. Users can fill entries to add new data, or they can add certain entries and DBMS will retrieve matching data for the rest.

Graphical User Interfaces:

It displays schema in a diagrammatic form. Most GUIs use pointing devices such as mouse to pick some parts of the schema.

Natural Language Interfaces:

It has its own schema. These interfaces accept requests written in natural languages and attempt to understand them. It generates a high level query corresponds to the natural language request and submits to the DBMS for processing.

Interfaces for Parametric Users:

Parametric users often have some set of operations that they must perform repeatedly. Programmers implement a special type of interface with a set of commands with the goal of reducing the number of key strokes for each request.

Interfaces for DBA:

Certain privileged commands such as commands for account creation, setting parameters of the system, granting the authorization, changing the schema, and reorganizing database structure can be used by DBA staff only.

1.2.4 Database system environment

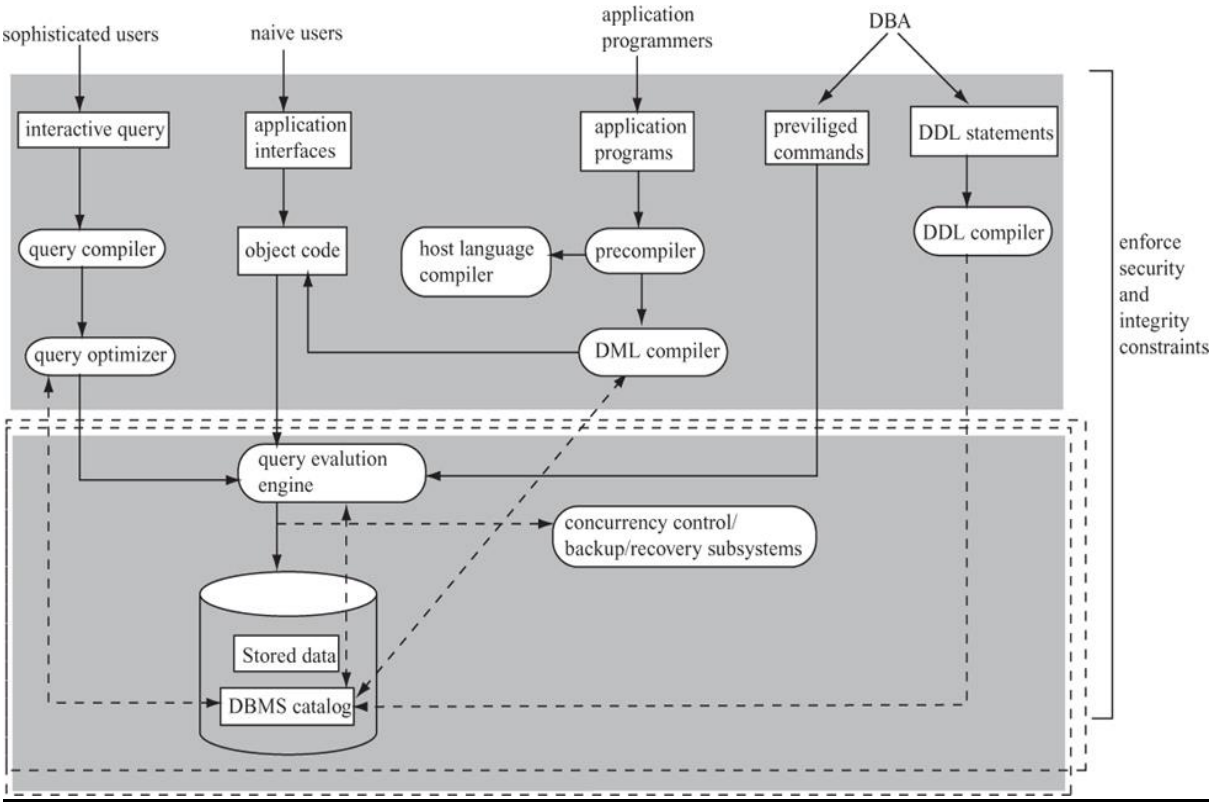


Figure1.6 DBMS Component Modules

The above **Figure 1.6** illustrates DBMS components. The database and catalog is stored in a disk. Operating System controls the accesses to the disk and it schedules the disk input /output. **Stored data manager** module of DBMS controls access to the disk information. It may use basic OS services for carrying out data transfer between storage and disk. **Buffer manager module** is using for buffering disk pages. **DDL compiler** processes all schema definitions and will store it in the catalog.

The **runtime database processor** controls database accesses at runtime. The **Query compiler** controls high level queries that are entered simultaneously. It analyses, compiles query by creating an access code and generates calls to the processor for execution. The **precompiler** extracts DML commands written in a programming language .These commands will be send to the **DML compiler** for converting to the object code. The client program resides in the client computer and DBMS resides in the database server.

Database System Utilities:

It helps DBA to manage the database system.

Loading Utility:

It is used to load some existing files such as text or sequential files in to db. The current format and the desired format are specified to the utility which automatically reformats data and will store in to database.

Backup Utility:

It creates backup copy of the database which helps to restore the database in case of catastrophic failures.

File reorganization:

It helps to reorganize the database to improve performance.

Performance monitoring:

It helps to monitor the performance. It monitors the database performance and provides statistics to DBA.

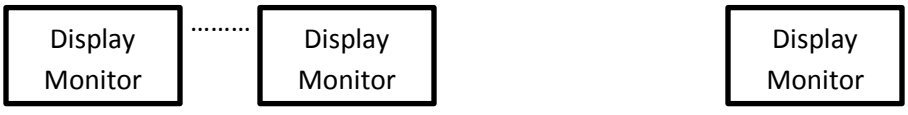
Tools Application Environments and Communication Facilities:

Data dictionary is an important tool using in large organizations. It stores application program descriptions, design decisions, usage standards, and other user information. Application Development Environments such as power builder system provides an environment to develop database applications. The integrated database management system and data communication system is called DB/DC system.

1.2.5. Centralized and client /server architecture for DBMS

Centralized DBMS Architecture:

Earlier architectures used mainframe computers for processing. Users accessed the systems via terminals and it has lack of processing power. All processing was at the remote computer. The display information will be sent to the terminals. The following figure shows



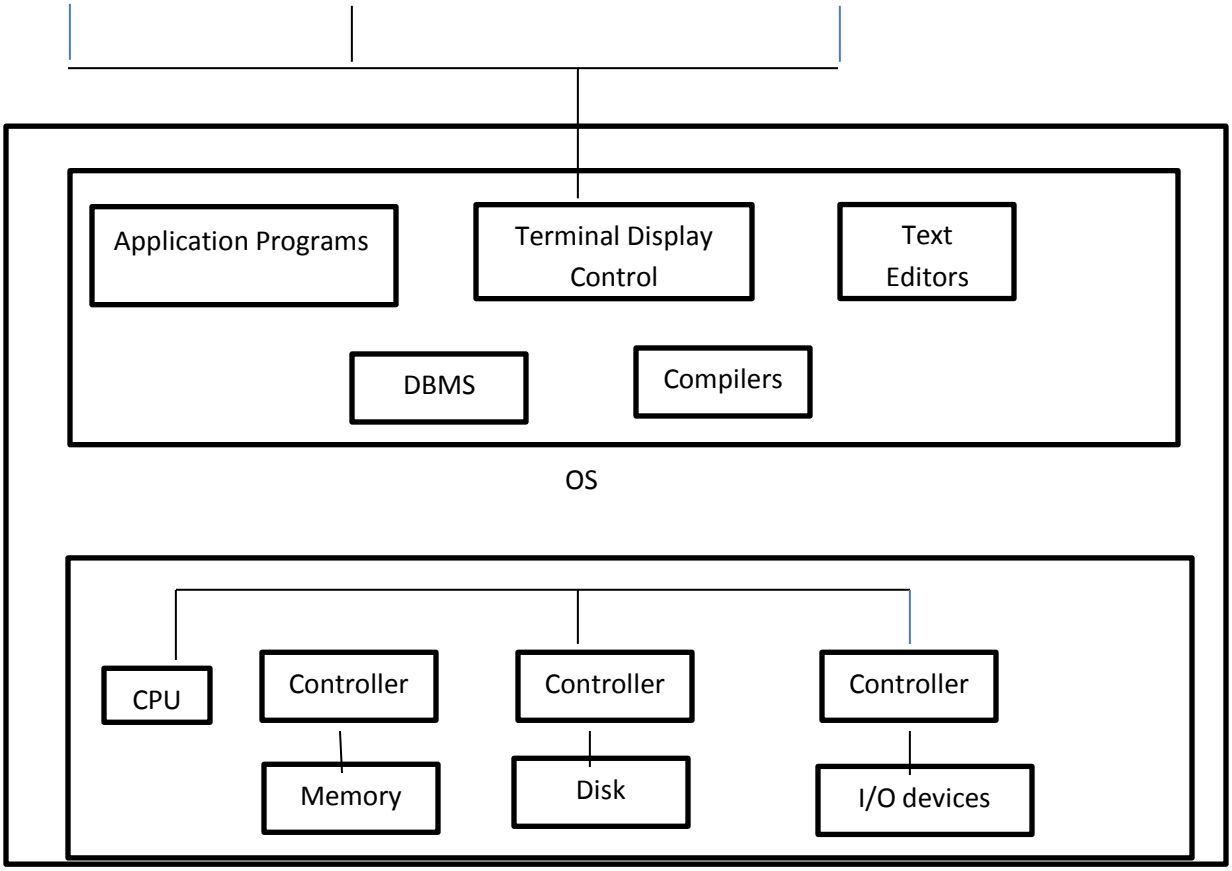


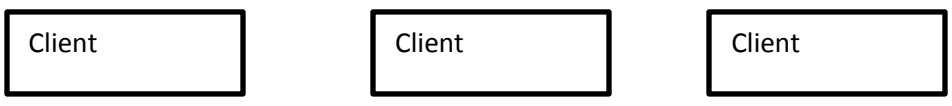
Figure1.7

Basic Client /Server Architectures:

- Specialized Servers with Specialized functions
- Print server
- File server
- DBMS server
- Web server
- Email server
- Clients are able to access the specialized servers as needed

Logical two-tier client server architecture:

Logical two-tier client server architecture is shown in figure 1.8



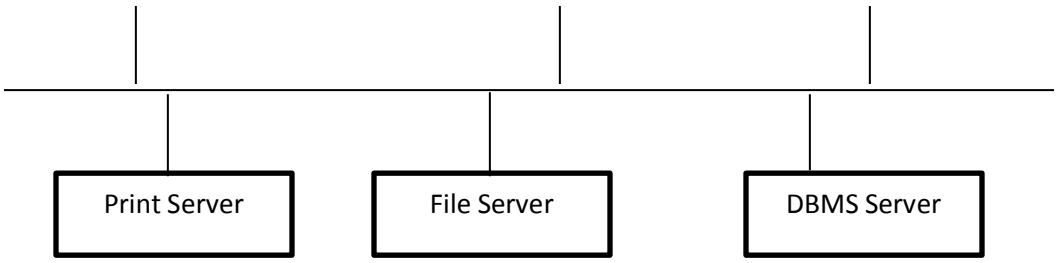


Figure1.8

Clients:

• Clients Offer proper interfaces through a client software module to access the server resources.

• Clients are connected to the servers with some kinds of networks such as LAN or wireless networks etc.

DBMS Server:

•It Provides queries as well as the transaction services to clients.

• Relational DBMS servers are also called query servers, SQL servers/transaction servers.

• Applications running on clients use an API to access server databases via standard interface.

Two tier Client server architecture for DBMS:

a) Client program connect many DBMSs which also called the data sources.

Three-Tier client/Server architecture for web applications:

• It is Common for the Web based applications.

• Intermediate Layer entitled as Application Server / Web Server.

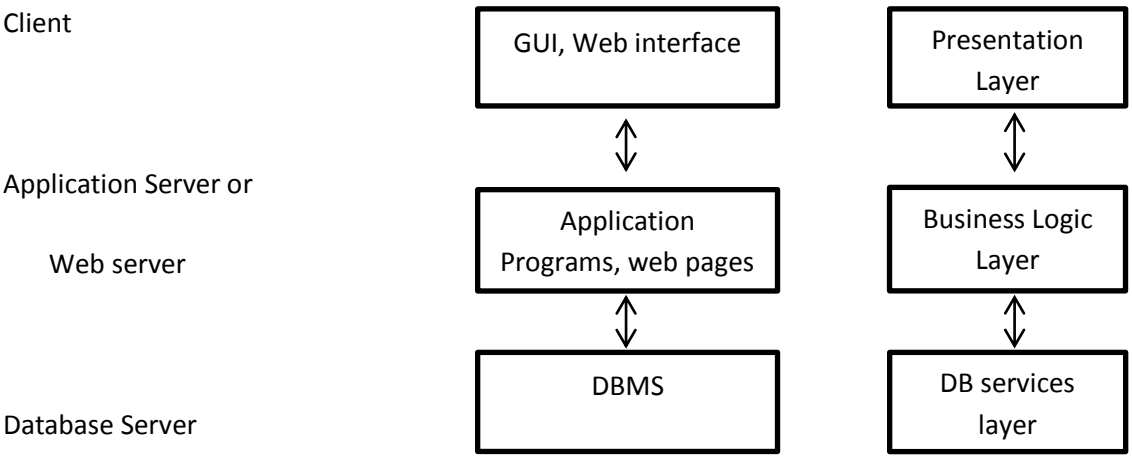
• Helps to Store the web connectivity software and the business logic part of the application to access the data from the database server.

• Acts like a channel for sending processed data between the database server and the client.

•It enhances security.

• Clients are not able to directly access database server.

Three-Tier client/Server architecture is shown in the following figure



Figue1.9 (a) (b)

1.2.6. Classification of database management systems

We can classify the DBMSs based on the data model on which it is based, number of users supported by the system, number of sites over which it is distributed and the types of access paths.

Based on the data models

We can classify the DBMS system as relational, object relational, object, network and hierarchical. The main data model using in many commercial DBMS s is the relational data model. Object data model is implemented in some commercial systems and many applications are still running in systems based on hierarchical and network models. Object relational DBMSs uses the concepts from relational databases as well as object databases. The relational model represents database as collection of tables and each table is stored as separate file. Object model represents database in terms of objects, properties and heir corresponding operations. The network model represents data as record types and represents type of 1:N relationship called set type. Hierarchical model represents data as tree structures.

Based on the number of users

We can classify system as Single user systems and multi user systems. Single user systems supports only one user at a time and multiuser system supports multiple users concurrently.

Based on number of sites

We can classify as centralized and distributed DBMS. DBMS is centralized when the data is stored at a single site. In distributed DBMS the database is distributed over multiple sites via networks.

Based on cost:

The cost ranges from 10K$-100K$ in case of many DBMS packages.

Based on Types of access paths:

DBMS can be general purpose or special purpose DBMS based on types of access paths. Special purpose DBMS s designed for improved performance. E.g.: Airline reservation system, telephone directory systems. They are coming under the category of Online Transaction Processing systems, which support large number of concurrent transactions.