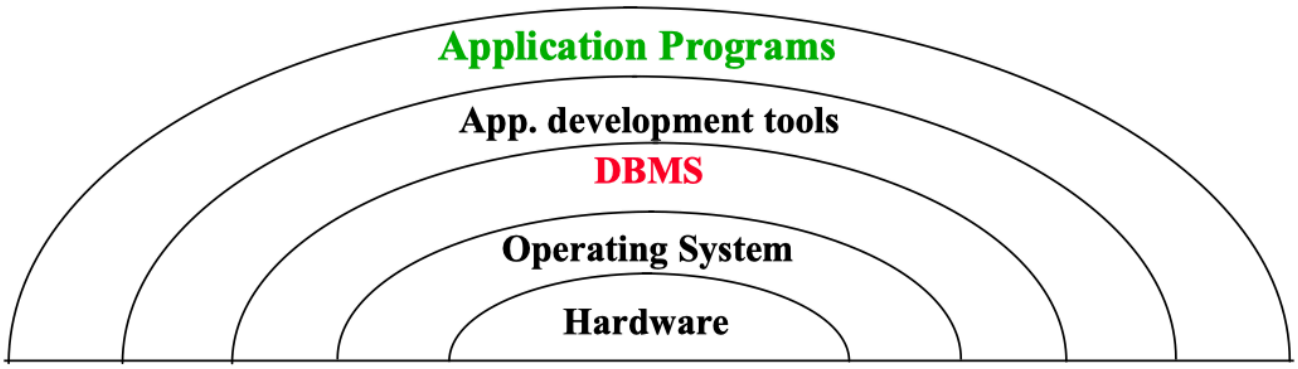
Overview

- A database is a very large, integrated, collection of data
- Models real-world enterprise

Database Management System (DBMS)

A DBMS is a software package designed to store and manage database



Basic Idea

1. Remove details related to data storage
2. Centralize system of the functions
3. Have all applications access data through the DBMS

Advantages

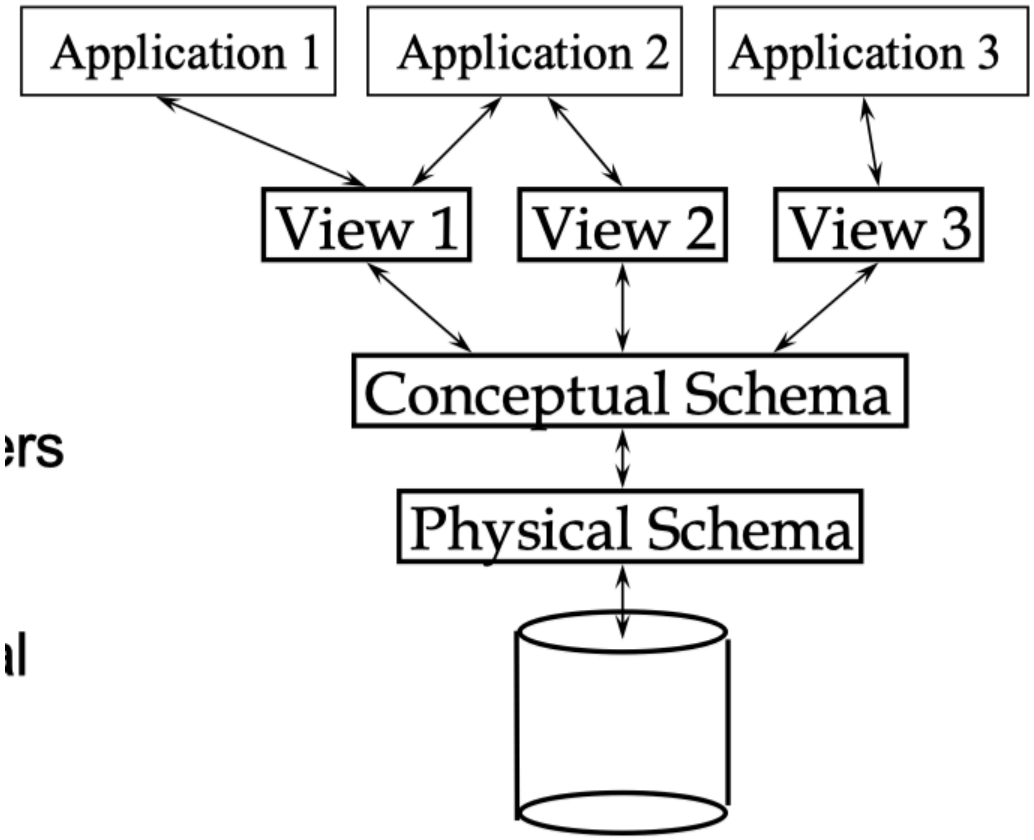
1. Reduced redundancy
2. Less risk of inconsistency
3. Reduced application development time
4. Uniform data administration
5. Concurrent access, recovery from crashes
6. Data Independence

Problems with Non-Database Approach

1. many files with different structures

2. Redundant storage
3. Inconsistent copies
4. Expensive updates
5. Incorrect data
6. Data exchange between applications

Three Schema Levels



A **Schema** is a description of the data contents, structures, and other aspects.

1. **External Schema**: What application programs and users see
   - the way data is seen by individual users
2. **Conceptual Schema**: Description of the logical structure of data
   - views database as a collection of tables in the relational model
3. **Physical Schema**: File structures and indexes being used
   - the way data is stored inside the system

Data Independence

The objective is application programs unaffected by changes in storage structure and access strategy

Logical Data Independence

Protection from changes in *logical* structure of data

Physical Data Independence

Protection from changes in *physical* structure of data

DBMS Functionality

1. **Data definition**: stores the definitions in a user-accessible catalogue
2. **Data Manipulation**: provides a query language for storing, retrieving, and updating data
3. **Integrity Constraints**: validation checks
4. **Concurrency Control**: allows for multiple user access
5. **Transactions**: sequence of operations that can be performed as an atomic action (all or nothing)
6. **Database Recovery**
7. **Query Optimization**