CS235/CS334– DATABASE TECHNOLOGY

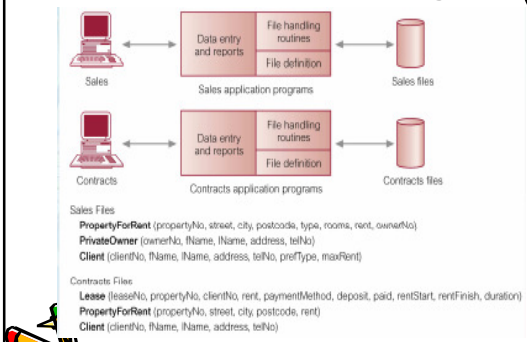
- CA – 40%
  - 3 TESTS
- EXAM – 60%
- www.LeChaamwe.weebly.com
  - Lecture Notes
  - Undergraduate
  - CS235 and CS334

Introduction to Databases

File-based Systems

- File-based systems were an early attempt to computerize the manual filing system
- File-based Systems defn
  - Collection of application programs that perform
  - services for the end users (e.g. reports).
  - Each program defines and manages its own data.

File-based Processing



Limitations of File-based Approach

- Separation and isolation of data
  - Each program maintains its own set of data.
  - Users of one program may be unaware of potentially useful data held by other programs.
- Duplication of data
  - Same data is held by different programs. Wasted space and potentially different values and/or different formats for the same item

Limitations of File-based Approach

- Data dependence
  - File structure is defined in the program code.
- Incompatible file formats
  - Programs are written in different languages, and so cannot easily access each others files.

Limitations of File-based Approach

- Fixed Queries/Proliferation of application programs
  - Programs are written to satisfy particular functions.
  - Any new requirement needs a new program.
- No control over access and manipulation of data beyond that imposed by application programs

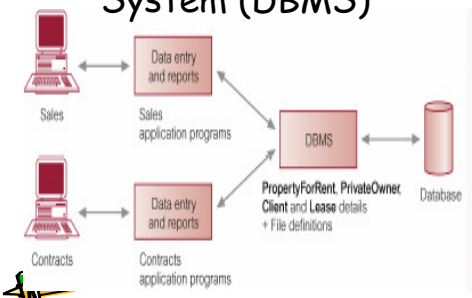
Database various definitions

- DB:
  - Shared collection of logically related data (and a description of this data), designed to meet the information needs of an organization.
  - System catalog (metadata) provides description of data to enable program-data independence.
  - Logically related data comprises entities, attributes, and relationships of an organization's information.

Database various definitions

- DBMS:
  - A software system that enables users to define, create, and maintain the database and which provides controlled access to this database.

Database Management System (DBMS)



Database architecture

- Three Level Architecture also known as ANSI-SPARC architecture, named after the committee that proposed it, the American National Standard Institute, Standards Planning And Requirements Committee.

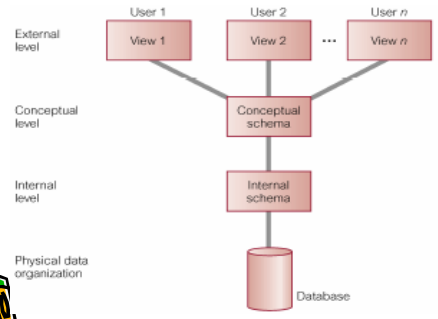
Objectives of Three-Level Architecture

- All users should be able to access same data.
- A user's view is immune to changes made in other views.
- Users should not need to know physical database storage details.

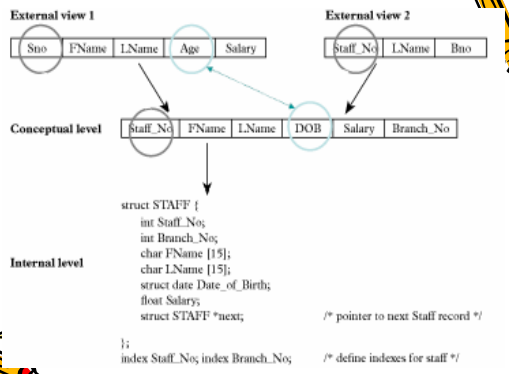
Objectives of Three-Level Architecture

- DBA should be able to change database storage structures without affecting the users' views.
- Internal structure of database should be unaffected by changes to physical aspects of storage.
- DBA should be able to change conceptual structure of database without affecting all

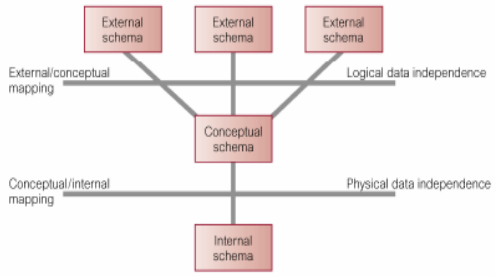
ANSI-SPARC Architecture



ANSISPARC Architecture



Data Independence and the 3 Levels of ANSISPARC Architecture



Data Independence

- Logical Data Independence
  - Refers to immunity of external schemas to changes in conceptual schema.
  - Conceptual schema changes (e.g. addition/removal of entities).
    - Should not require changes to external schema or rewrites of application programs.

Data Independence

- Physical Data Independence
  - Refers to immunity of conceptual schema to changes in the internal schema.
  - Internal schema changes (e.g. using different file organizations, storage structures/devices).
  - Should not require change to conceptual or external schemas.

Functions of a DBMS

- Data Storage, Retrieval, and Update.
- A User-Accessible Catalog.
- Transaction Support.
- Concurrency Control Services.
- Recovery Services.

Functions of a DBMS

- Authorization Services.
- Support for Data Communication.
- Integrity Services.
- Services to Promote Data Independence.
- Utility Services.

Advantages of DBMS

- Control of data redundancy
- Data consistency
- More information from the same amount of data.
- Sharing of data
- Improved data integrity
- Improved security
- Enforcement of standards
- Economy of scale

Advantages of DBMS

- Balanced conflicting requirements
- Improved data accessibility and responsiveness
- Increased productivity
- Improved maintenance through data independence
- Increased concurrency
- Improved backup and recovery services

Disadvantages of DBMS

- Complexity
- Size
- Cost of DBMS
- Additional hardware costs
- Cost of conversion
- Performance
- Higher impact of a failure

Adv and Disadv of DBMS

- For details of these advantages and disadvantages, read the attached copy on the website.

When a DBMS may be unnecessary

- If the database and applications are simple, well defined, and not expected to change
- If there are stringent real-time requirements that may not be met because of DBMS overhead.
- If access to data by multiple users is not required.

Questions