Introduction to Databases
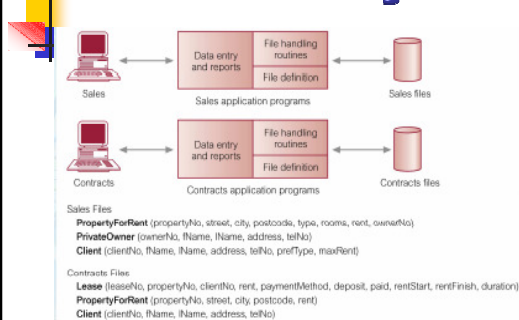
Dr N Chaamwe

Database systems vs. file-based systems

- File-based Systems
  - Collection of application programs that perform
  - services for the end users (e.g. reports).
  - Each program defines and manages its own data.

Dr N Chaamwe

File-based Processing

Sales Files
PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo).
PrivateOwner (ownerNo, fName, lName, address, telNo).
Client (clientNo, fName, lName, address, telNo, prefType, maxRent).

Contracts Files
Lease (leaseNo, propertyNo, clientNo, rent, paymentMethod, deposit, paid, rentStart, rentFinish, duration).
PropertyForRent (propertyNo, street, city, postcode, rent).
Client (clientNo, fName, lName, address, telNo).

Dr N Chaamwe

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

Dr N Chaamwe

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.
- Fixed Queries/Proliferation of application programs
  - Programs are written to satisfy particular functions.
  - Any new requirement needs a new program.

Dr N Chaamwe

Database Approach

- Arose because:
  - Definition of data was embedded in application programs, rather than being stored separately and independently.
  - No control over access and manipulation of data beyond that imposed by application programs.
- Result
  - the database and Database Management System (DBMS). Dr N Chaamwe

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.
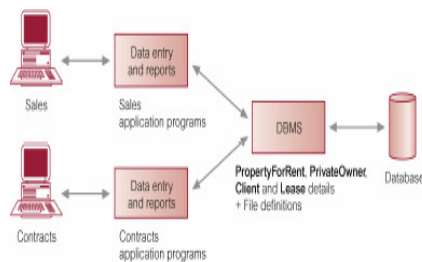
Dr N Chaamwe

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.

Dr N Chaamwe

Database Management System (DBMS)

Dr N Chaamwe

Three-level schema architecture

- ANSI-SPARC architecture is named after the committee that proposed it, the American National Standard Institute, Standards Planning And Requirements Committee.

Dr N Chaamwe

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.

Dr N Chaamwe

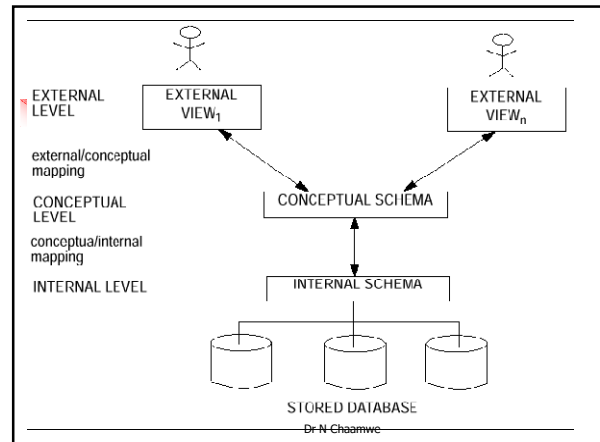## Objectives of Three-Level Architecture

- 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.

Dr N Chaamwe

Objectives of Three-Level Architecture

- DBA should be able to change conceptual structure of database without affecting all users.

Dr N Chaamwe

EXTERNAL LEVEL — EXTERNAL VIEW$_1$ ... EXTERNAL VIEW$_n$

external/conceptual mapping

CONCEPTUAL LEVEL — CONCEPTUAL SCHEMA

conceptual/internal mapping

INTERNAL LEVEL — INTERNAL SCHEMA

STORED DATABASE

Dr N Chaamwe

The External level

- The external level (also known as the user logical level) is the one closest to the users.
- It is the level concerned with the way the data is seen by individual users.

Dr N Chaamwe

The External level

- The external schema is used to describe the external level.
- It contains a description of a portion of the database that is of concern to the specific user.

Dr N Chaamwe

The External level

- The external view is described in terms of external records, which may be different form the actual stored records.
- The view may have a different representation of the same data

Dr N Chaamwe

The Conceptual level

- The level represents the community view of the database as seen by the database administrator.
- It is also known as the community logical level or even sometimes just the logical level.

Dr N Chaamwe

The Conceptual level

- The conceptual schema is used to describe what data is stored in the database and the relationships among the data.
- The description includes the structure and constraints for the whole database

Dr N Chaamwe

The Conceptual level

- the conceptual schema defines the logical structure of all data in the database.
- The conceptual schema is defined by a Data Definition Language (DDL).
- There is only one conceptual schema for the database.

Dr N Chaamwe

The Internal level

- It is the description of the implementation of the conceptual schema by means of physical storage structures.
- It summarizes how the data are stored on secondary storage devices such as disks and tapes.

Dr N Chaamwe

The Internal level

- However, it is still one level above the actual physical storage, which is usually managed by the operating system.
- The internal level does not deal with the physical records but it deals with the internal records (stored records).

Dr N Chaamwe

The Internal level

- Details of how the address space is mapped to physical storage are highly system-specific (e.g. a block or a page)

Dr N Chaamwe

External view 1						External view 2		
Sno	FName	LName	Age	Salary		Staff_No	LName	Bno

Conceptual level	Staff_No	FName	LName	DOB	Salary	Branch_No

Dr N Chaamwe

```
                              '
        struct STAFF {
              int Staff_No;
              int Branch_No;
              char FName [15];
              char LName [15];
              struct date Date_of_Birth;
              float Salary;
              struct STAFF *next;          /* pointer to next Staff rec

internal level
        };
        index Staff_No; index Branch_No;   /* define indexes for staff
```
Dr N Chaamwe

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.

Dr N Chaamwe

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.

Dr N Chaamwe

Data Independence

- Describes the immunity of the upper levels from the changes in the lower levels

Dr N Chaamwe

Mappings

- The process of transforming requests and results between levels are called mappings.
- The three-level architecture involves certain mappings— one conceptual/internal mapping and several external/conceptual mappings.

Dr N Chaamwe

The conceptual/internal mapping

- defines the correspondence between the conceptual view and the stored databases
- If a change is made to the storage definition, then the conceptual/internal mapping must be changed accordingly, so that the conceptual schema can remain invariant
Dr N Chaamwe

The external/conceptual mapping

- defines the correspondence between a particular external view and the conceptual view.
- If a change is made to the conceptual schema definition, then the External/conceptual mapping must be changed accordingly, so that the External Schema can remain invariant.

Dr N Chaamwe

Functions of a DBMS

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

Dr N Chaamwe

Functions of a DBMS

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

Dr N Chaamwe

Advantages of DBMS

- Control of data redundancy
- Data consistency
- More information from the same amount of data.
- Sharing of data

Dr N Chaamwe

Advantages of DBMS

- Improved data integrity
- Improved security
- Enforcement of standards
- Economy of scale
- Improved data accessibility and responsiveness

Dr N Chaamwe

Advantages of DBMS

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

Dr N Chaamwe

Disadvantages of DBMS

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

Dr N Chaamwe

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.

Dr N Chaamwe