Introduction to Database

Lecture 01

Dr. Reda M. Hussien

File Terminology

- Data
  - Raw Facts
- Field
  - Group of characters with specific meaning
- Record
  - Logically connected fields that describe a person, place, or thing
- File
  - Collection of related records

Lease

leaseNo	propertyNo	clientNo	rent	payment Method	deposit	paid	rentStart	rentFinish	duration
10024	PA14	CR62	650	Visa	1300	Y	1-Jun-13	31-May-14	12
10075	PL94	CR76	400	Cash	800	N	1-Aug-13	31-Jan-14	6
10012	PG21	CR74	600	Cheque	1200	Y	1-Jul-13	30-Jun-14	12

PropertyForRent

propertyNo	street	city	postcode	rent
PA14	16 Holhead	Aberdeen	AB7 5SU	650
PL94	6 Argyll St	London	NW2	400
PG21	18 Dale Rd	Glasgow	G12	600

Client

clientNo	fName	lName	address	telNo
CR76	John	Kay	56 High St, London SW1 4EH	0171-774-5632
CR74	Mike	Ritchie	18 Tain St, PA1G 1YQ	01475-392178
CR62	Mary	Tregear	5 Tarbot Rd, Aberdeen AB9 3ST	01224-196720

Traditional File-based Systems

- File-based System
  - A collection of application programs that perform services for the end-users such as the production of reports. Each program defines and manages its own data.

Simple File System

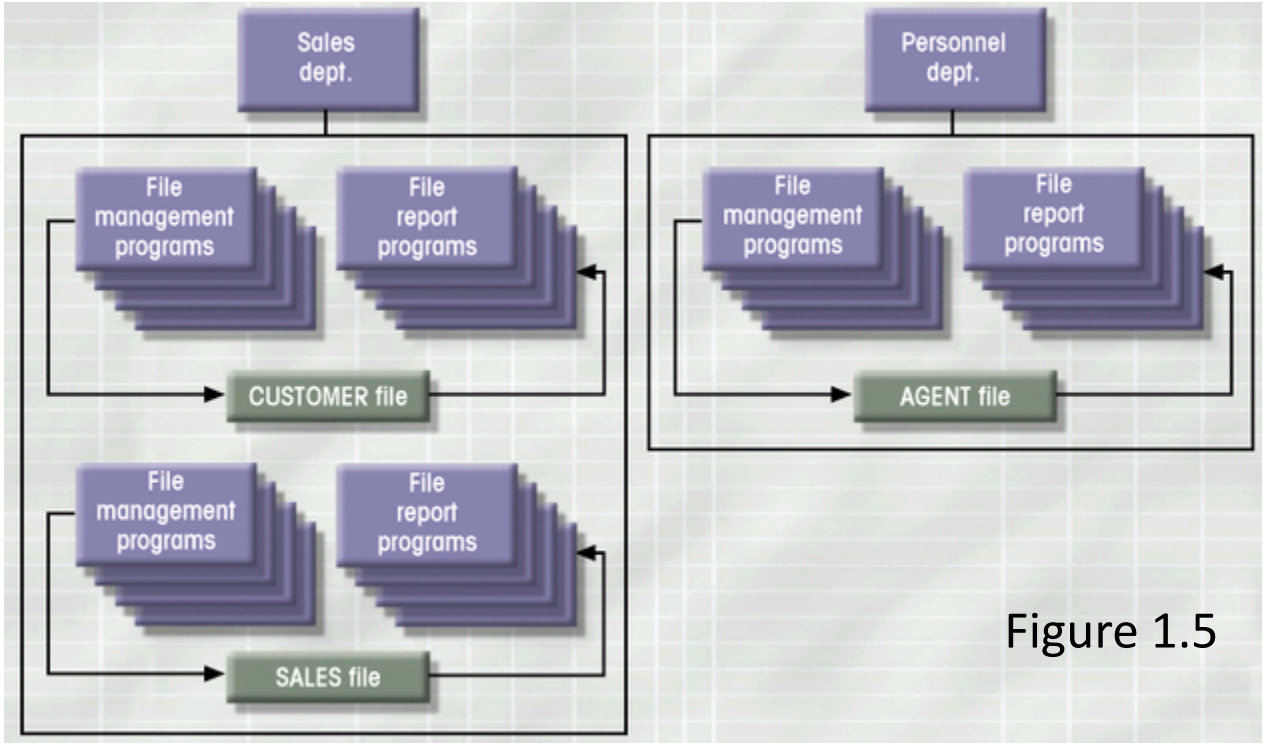


Figure 1.5

Traditional File-based Systems

- Information are stored in data files
- Each file is a sequence of records

Lease

leaseNo	propertyNo	clientNo	rent	payment Method	deposit	paid	rentStart	rentFinish	duration
10024	PA14	CR62	650	Visa	1300	Y	1-Jun-13	31-May-14	12
10075	PL94	CR76	400	Cash	800	N	1-Aug-13	31-Jan-14	6
10012	PG21	CR74	600	Cheque	1200	Y	1-Jul-13	30-Jun-14	12

PropertyForRent

propertyNo	street	city	postcode	rent
PA14	16 Holhead	Aberdeen	AB7 5SU	650
PL94	6 Argyll St	London	NW2	400
PG21	18 Dale Rd	Glasgow	G12	600

Client

clientNo	fName	lName	address	telNo
CR76	John	Kay	56 High St, London SW1 4EH	0171-774-5632
CR74	Mike	Ritchie	18 Tain St, PA1G 1YQ	01475-392178
CR62	Mary	Tregear	5 Tarbot Rd, Aberdeen AB9 3ST	01224-196720

Traditional File-based Systems

- Eg., if each record contains 100 bytes of data, then
- the 1st record occupies the 1st 100 bytes in the file,
- the 2nd record occupies the 2nd 100 bytes in the file.

Lease

leaseNo	propertyNo	clientNo	rent	payment Method	deposit	paid	rentStart	rentFinish	duration
10024	PA14	CR62	650	Visa	1300	Y	1-Jun-13	31-May-14	12
10075	PL94	CR76	400	Cash	800	N	1-Aug-13	31-Jan-14	6
10012	PG21	CR74	600	Cheque	1200	Y	1-Jul-13	30-Jun-14	12

PropertyForRent

propertyNo	street	city	postcode	rent
PA14	16 Holhead	Aberdeen	AB7 5SU	650
PL94	6 Argyll St	London	NW2	400
PG21	18 Dale Rd	Glasgow	G12	600

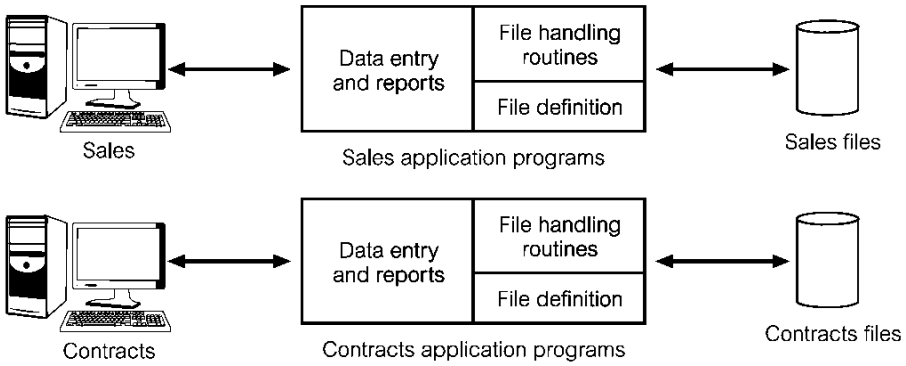
Client

clientNo	fName	lName	address	telNo
CR76	John	Kay	56 High St, London SW1 4EH	0171-774-5632
CR74	Mike	Ritchie	18 Tain St, PA1G 1YQ	01475-392178
CR62	Mary	Tregear	5 Tarbot Rd, Aberdeen AB9 3ST	01224-196720

Traditional File-based Systems

- Limitations of File-Based Approach:
  - Separation and isolation of data
  - Duplication of data
  - Data dependence
  - Incompatibility of files
  - Fixed queries / proliferation of application programs

Traditional File-based Systems



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)

Lease

leaseNo	propertyNo	clientNo	rent	payment Method	deposit	paid	rentStart	rentFinish	duration
10024	PA14	CR62	650	Visa	1300	Y	1-Jun-13	31-May-14	12
10075	PL94	CR76	400	Cash	800	N	1-Aug-13	31-Jan-14	6
10012	PG21	CR74	600	Cheque	1200	Y	1-Jul-13	30-Jun-14	12

PropertyForRent

propertyNo	street	city	postcode	rent
PA14	16 Holhead	Aberdeen	AB7 5SU	650
PL94	6 Argyll St	London	NW2	400
PG21	18 Dale Rd	Glasgow	G12	600

Client

clientNo	fName	lName	address	telNo
CR76	John	Kay	56 High St, London SW1 4EH	0171-774-5632
CR74	Mike	Ritchie	18 Tain St, PA1G 1YQ	01475-392178
CR62	Mary	Tregear	5 Tarbot Rd, Aberdeen AB9 3ST	01224-196720

Traditional File-based Systems

- duplication of data is undesirable for several reasons, including:
  - Duplication is wasteful. It costs time and money to enter the data.
  - It takes up additional storage space, again with associated costs.
  - duplication can lead to loss of data integrity; in other words, the data is no longer consistent.

Traditional File-based Systems

Lease

leaseNo	propertyNo	clientNo	rent	payment Method	deposit	paid	rentStart	rentFinish	duration
10024	PA14	CR62	650	Visa	1300	Y	1-Jun-13	31-May-14	12
10075	PL94	CR76	400	Cash	800	N	1-Aug-13	31-Jan-14	6
10012	PG21	CR74	600	Cheque	1200	Y	1-Jul-13	30-Jun-14	12

PropertyForRent

propertyNo	street	city	postcode	rent
PA14	16 Holhead	Aberdeen	AB7 5SU	650
PL94	6 Argyll St	London	NW2	400
PG21	18 Dale Rd	Glasgow	G12	600

Client

clientNo	fName	lName	address	telNo
CR76	John	Kay	56 High St, London SW1 4EH	0171-774-5632
CR74	Mike	Ritchie	18 Tain St, PA1G 1YQ	01475-392178
CR62	Mary	Tregear	5 Tarbot Rd, Aberdeen AB9 3ST	01224-196720

Traditional File-based Systems

- Data dependence
  - the physical structure and storage of the data files and records are defined in the application code. This means that changes to an existing structure are difficult to make.

Traditional File-based Systems

- Incompatibility of files
  - Because the structure of files is embedded in the application programs, the structures are dependent on the application programming language.
  - For example, the structure of a file generated by a COBOL program may be different from the structure of a file generated by a C program.

Traditional File-based Systems

- Fixed queries / proliferation of application programs
  - file-based systems are very dependent upon the application developer,
  - the type of query or report that could be produced was fixed.
  - There was no facility for asking unplanned queries.

Database Approach

- All of the previously mentioned limitations of the file-based approach can be attributed to two factors:
  - The definition of the data is embedded in the application programs, rather than being stored separately and independently.
  - There is no control over the access and manipulation of data beyond that imposed by the application programs.

Database Approach

- The Database
  - Database A shared collection of logically related data and its description designed to meet the information needs of an organization.

Database Approach

- The Database
  - The database is a single, possibly large repository of data that can be used simultaneously by many departments and users.
  - Instead of disconnected files with redundant data, all data items are integrated with a minimum amount of duplication.
  - The database is no longer owned by one department but is a shared corporate resource.

Database Approach

- The Database
  - The database holds not only the organization's operational data, but also a description of this data.
  - For this reason, a database is also defined as a self-describing collection of integrated records.
  - The description of the data is known as the system catalog (or data dictionary or metadata—the "data about data").

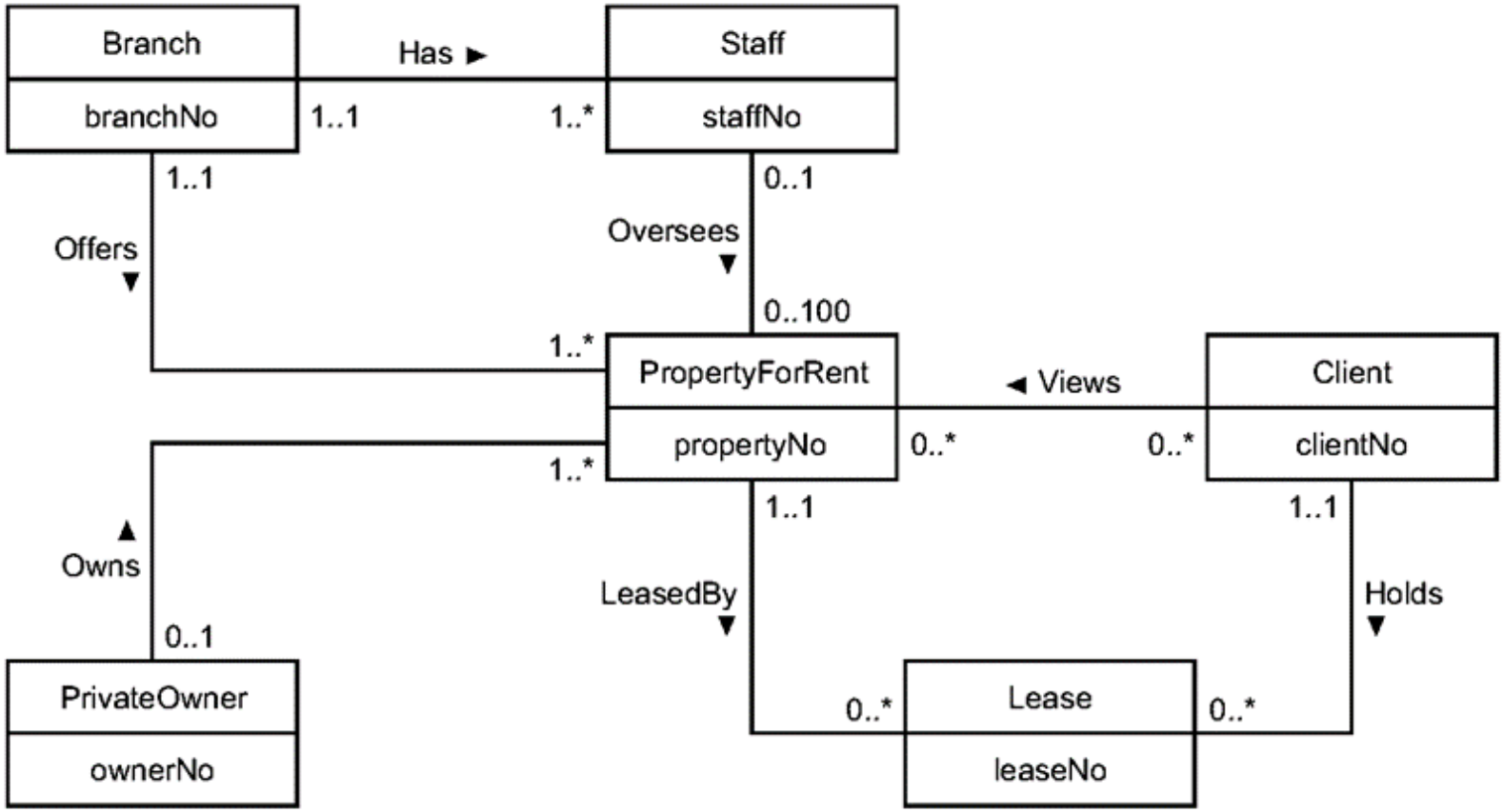
Database Approach

- The Database
  - the definition of data is separated from the application programs

Database Approach

- The Database Is logically related
  - An entity is a distinct object (a person, place, thing, concept, or event) in the organization that is to be represented in the database.
  - An attribute is a property that describes some aspect of the object that we wish to record, and
  - a relationship is an association between entities.

Database Approach



The Database Management System (DBMS)

- A software system that enables users to define, create, maintain, and control access to the database.
    - It allows users to define the database, usually through a Data Definition Language (DDL).
    - It allows users to insert, update, delete, and retrieve data from the database, usually through a Data Manipulation Language (DML).

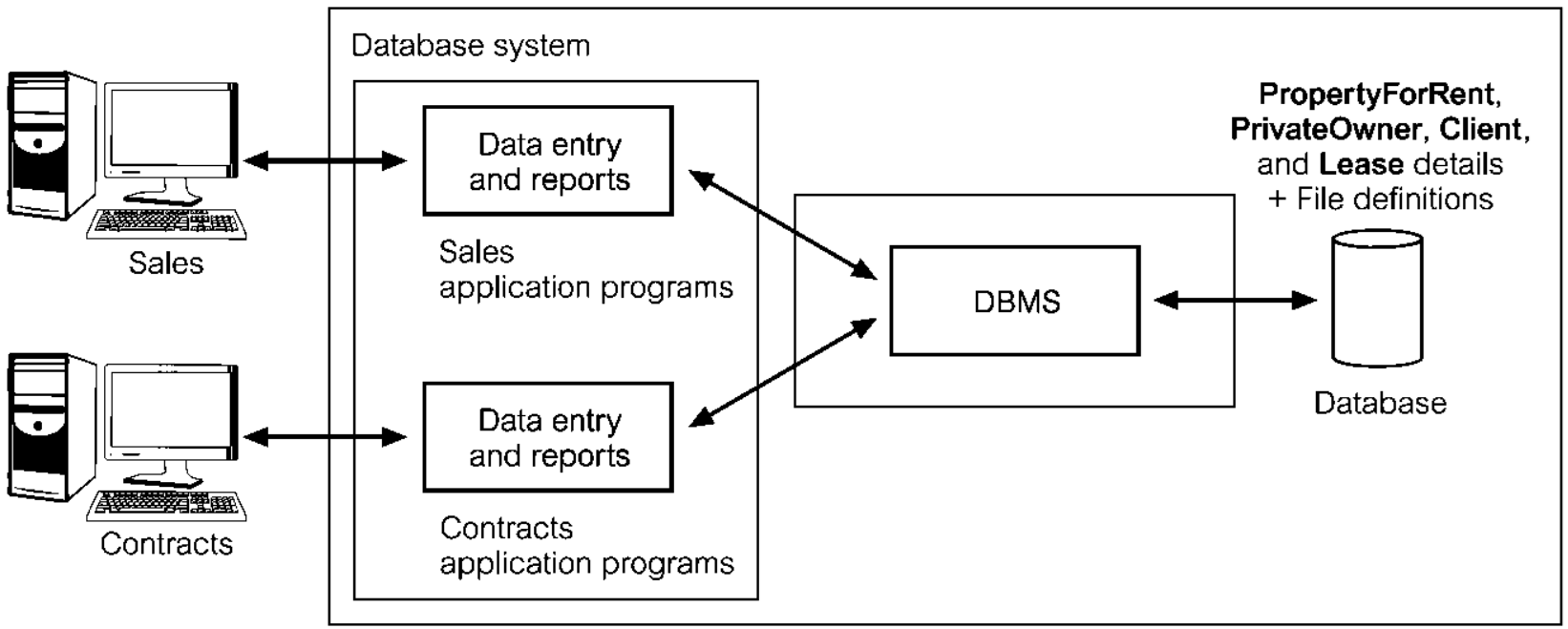
The Database Management System (DBMS)

- It provides controlled access to the database.
  - a security system, which prevents unauthorized users accessing the database;
  - an integrity system, which maintains the consistency of stored data;
  - a concurrency control system, which allows shared access of the database;
  - a recovery control system, which restores the database to a previous consistent state following a hardware or software failure;
  - a user-accessible catalog, which contains descriptions of the data in the database.

Application Programs

- Application Programs A computer program that interacts with the database by issuing an appropriate request (typically an SQL statement) to the DBMS.

Application Programs



PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)
PrivateOwner (ownerNo, fName, lName, address, telNo)
Client (clientNo, fName, lName, address, telNo, prefType, maxRent)
Lease (leaseNo, propertyNo, clientNo, paymentMethod, deposit, paid, rentStart, rentFinish)

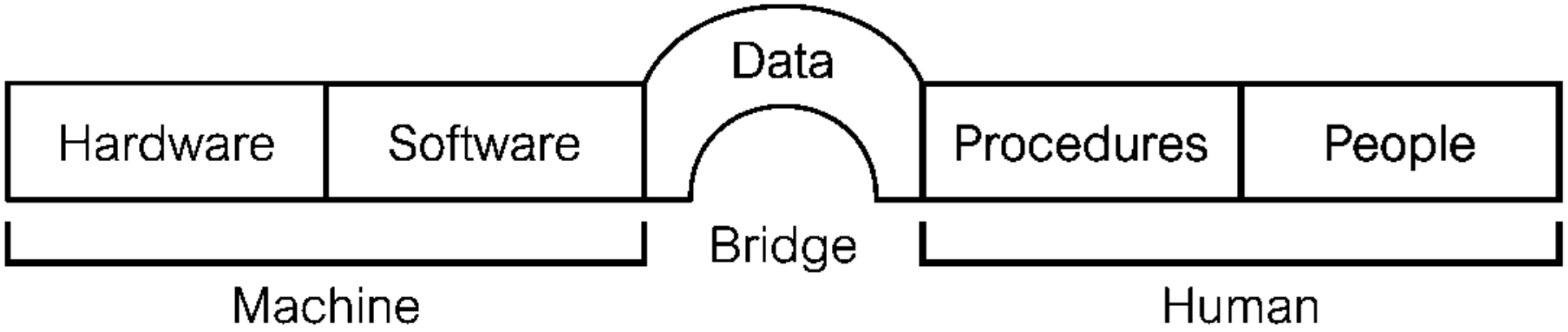
Views

- As well as reducing complexity by letting users see the data in the way they want to see it, views have several other benefits:

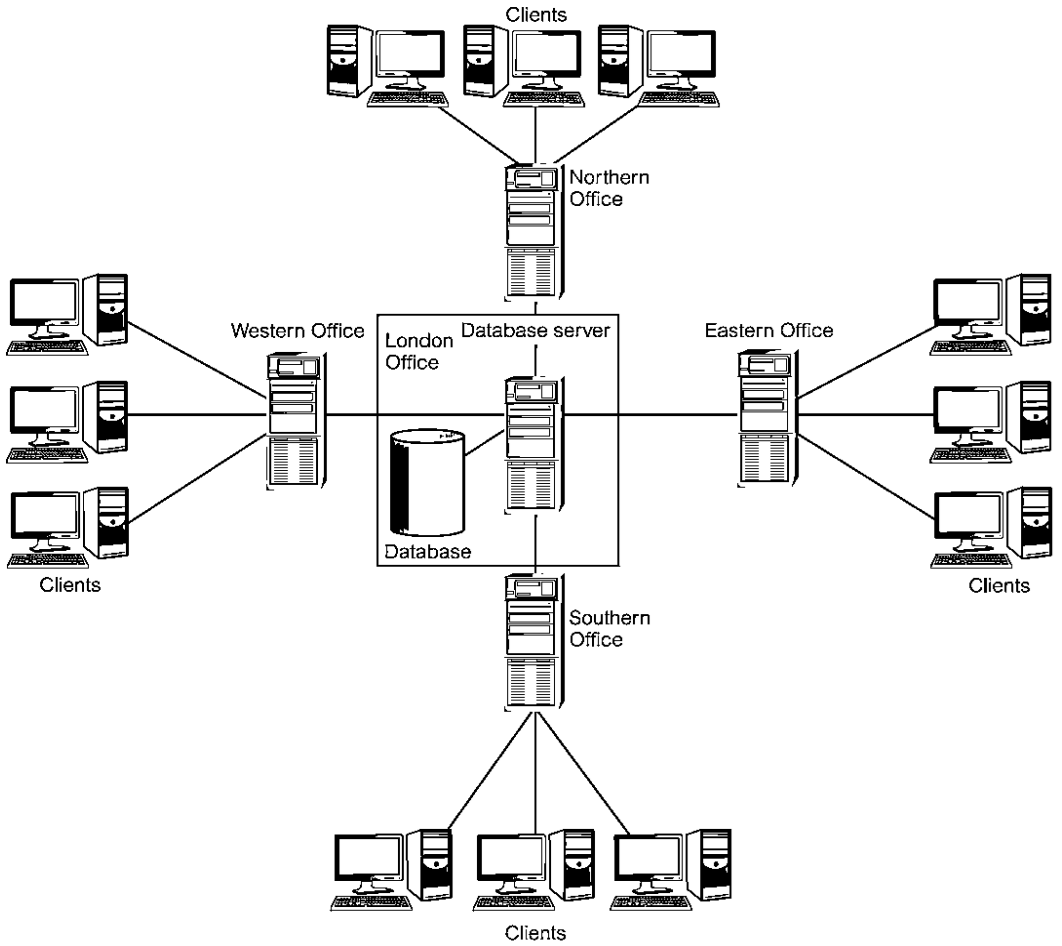
Views

- views have several other benefits
  - Views provide a level of security
  - Views provide a mechanism to customize the appearance of the database.
  - A view can present a consistent, unchanging picture of the structure of the database, even if the underlying database is changed

Components of the DBMS Environment



Components of the DBMS Environment



Components of the DBMS Environment

- Software
  - The software component comprises the DBMS software itself
  - the application programs
  - operating system, including network software if the DBMS is being used over a network.

Components of the DBMS Environment

- **Data**
  - Perhaps the most important component of the DBMS environment—certainly from the end-users' point of view

Components of the DBMS Environment

- **Procedures**
  - Procedures refer to the instructions and rules that govern the design and use of the database.
  - may consist of instructions on how to:
    - Log on to the DBMS.
    - Use a particular DBMS facility or application program.
    - Start and stop the DBMS.
    - Make backup copies of the database.
    - Handle hardware or software failures.

Components of the DBMS Environment

- **People**
  - The final component is the people involved with the system.
    - Data and Database Administrators
    - Database Designers
    - Application Developers
    - End-Users
      - Naïve users: unaware of the DBMS
      - Sophisticated users: familiar with the structure of the database.