Database Management systems

Data

Database

Collection of data organized in a manner that allows access, retrieval, and use of that data

Data

Collection of unprocessed items

- Text
- Numbers
- Images
- Audio
- Video

Information

Processed data

- Audio
- Documents
- Images
- Video

Database

- Stores a collection of related items
- Collection is arranged in a structure
  - Organizes and describes the data
- Often includes helper documents

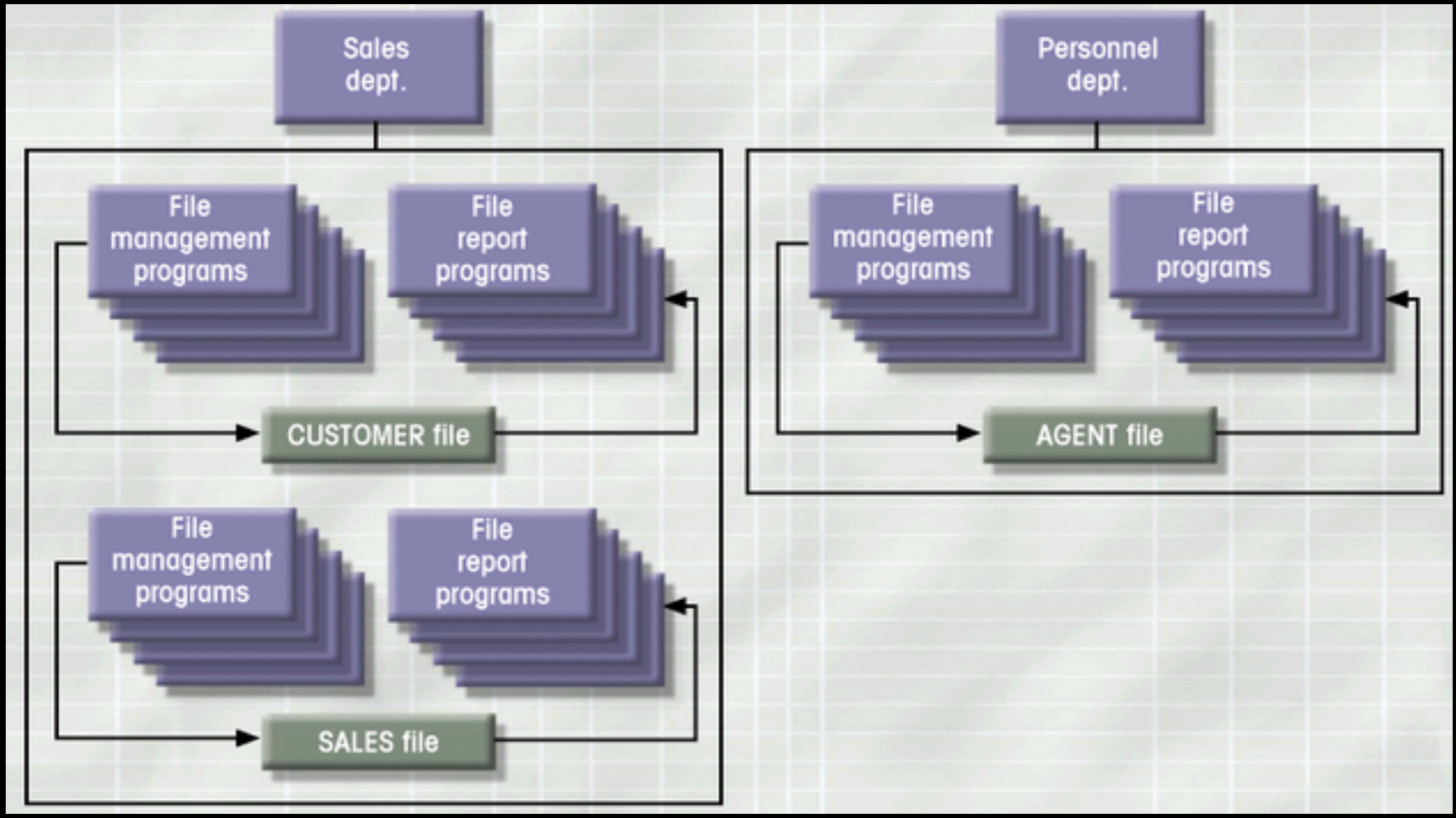
Database

- A database is a collection of information that is organized so that it can easily be accessed, managed, and updated. In one view, databases can be classified according to types of content: bibliographic, full-text, numeric, and images.

Database Applications

- Banking: all transactions
- Airlines: reservations, schedules
- Universities: registration, grades
- Sales: customers, products, purchases
- Manufacturing: production, inventory, orders, supply chain
- Human resources: employee records, salaries, tax deductions

File system


Sales dept.
Personnel dept.
File management programs
File report programs
CUSTOMER file
File management programs
File report programs
AGENT file
File management programs
File report programs
SALES file

For example in file

- The Payroll Department stores details relating to each member of staff's salary, namely:
- StaffSalary(staffNo, fName, lName, salary, branchNo)
- The Personnel Department also stores staff details, namely:
- Staff(staffNo, fName, lName, position, dateOfBirth, salary, branchNo)

Limitation of file base approach

- Separation and isolation of data
- Duplication of data
- 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.
  - For example, increasing the size of the PropertyForRent address field from 40 to 41 characters sounds like a simple change, but it requires the creation of a one-off program (that is, a program that is run only once and can then be discarded) that converts the PropertyForRent file to the new format.

Limitation of file base approach

- **Incompatible file formats**
- **Fixed queries/proliferation of application programs**
  - there was no provision for security or integrity
  - recovery, in the event of a hardware or software failure, was limited or non-existent;
  - access to the files was restricted to one user at a time – there was no provision for shared access by staff in the same department.

Database

- **Database** A shared collection of logically related data, and a description of this data, designed to meet the information needs of an organization.
- **DBMS** A software system that enables users to define, create, maintain, and control access to the database.

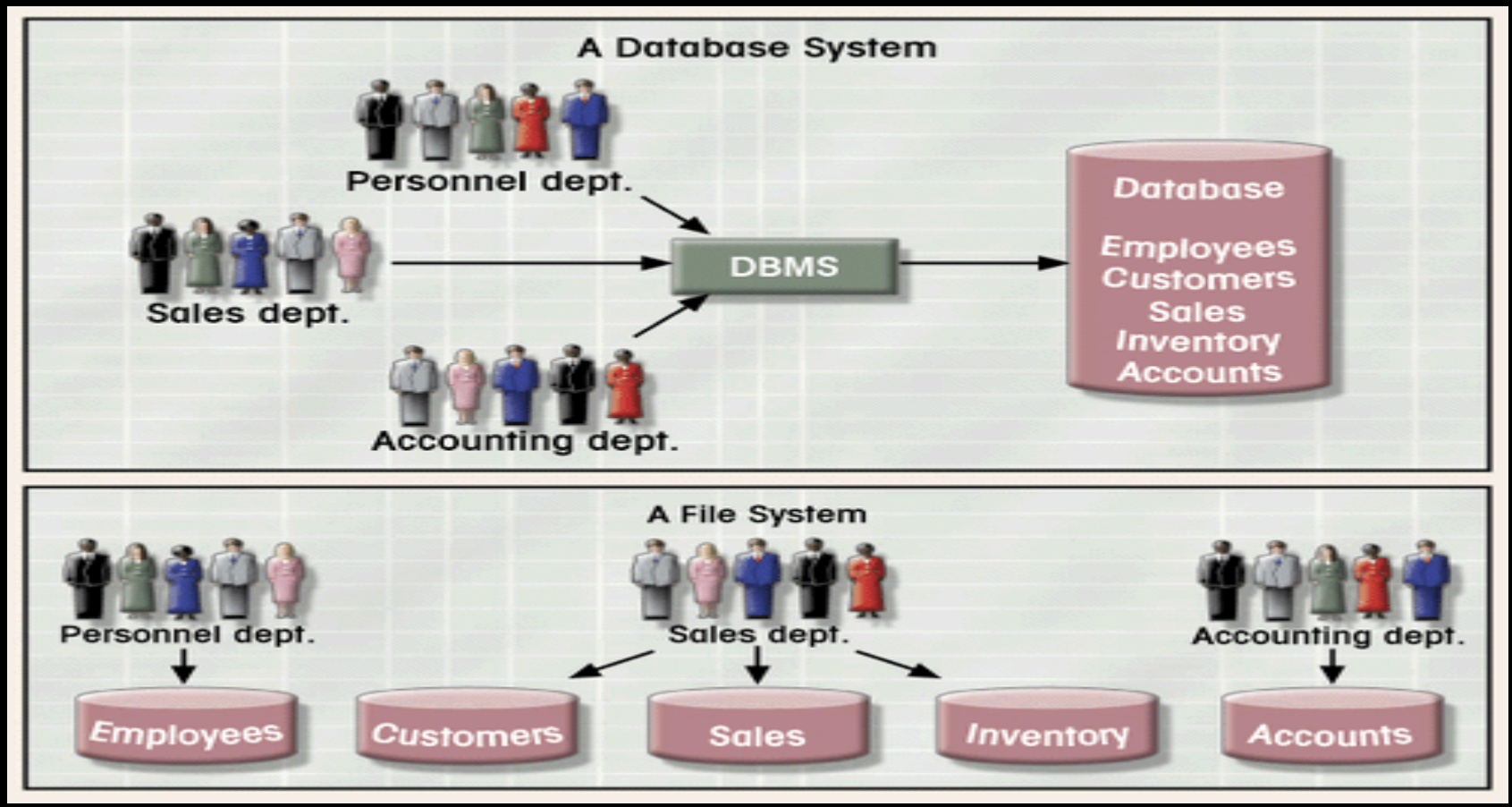
Purpose of Database System

- In the early days, database applications were built on top of file systems, can b used small-group situation
- Drawbacks of using file systems to store data:
  - Data redundancy and inconsistency(not same at different place)
    - Multiple file formats, duplication of information in different files
  - Difficulty in accessing data
    - Need to write a new program to carry out each new task
  - Data isolation — multiple files and formats(date format)
- Data dependence- This means that changes to an existing structure are difficult to make.
  - Integrity problems
    - Integrity constraints (e.g. account balance > 0) become part of program code
    - Hard to add new constraints or change existing ones

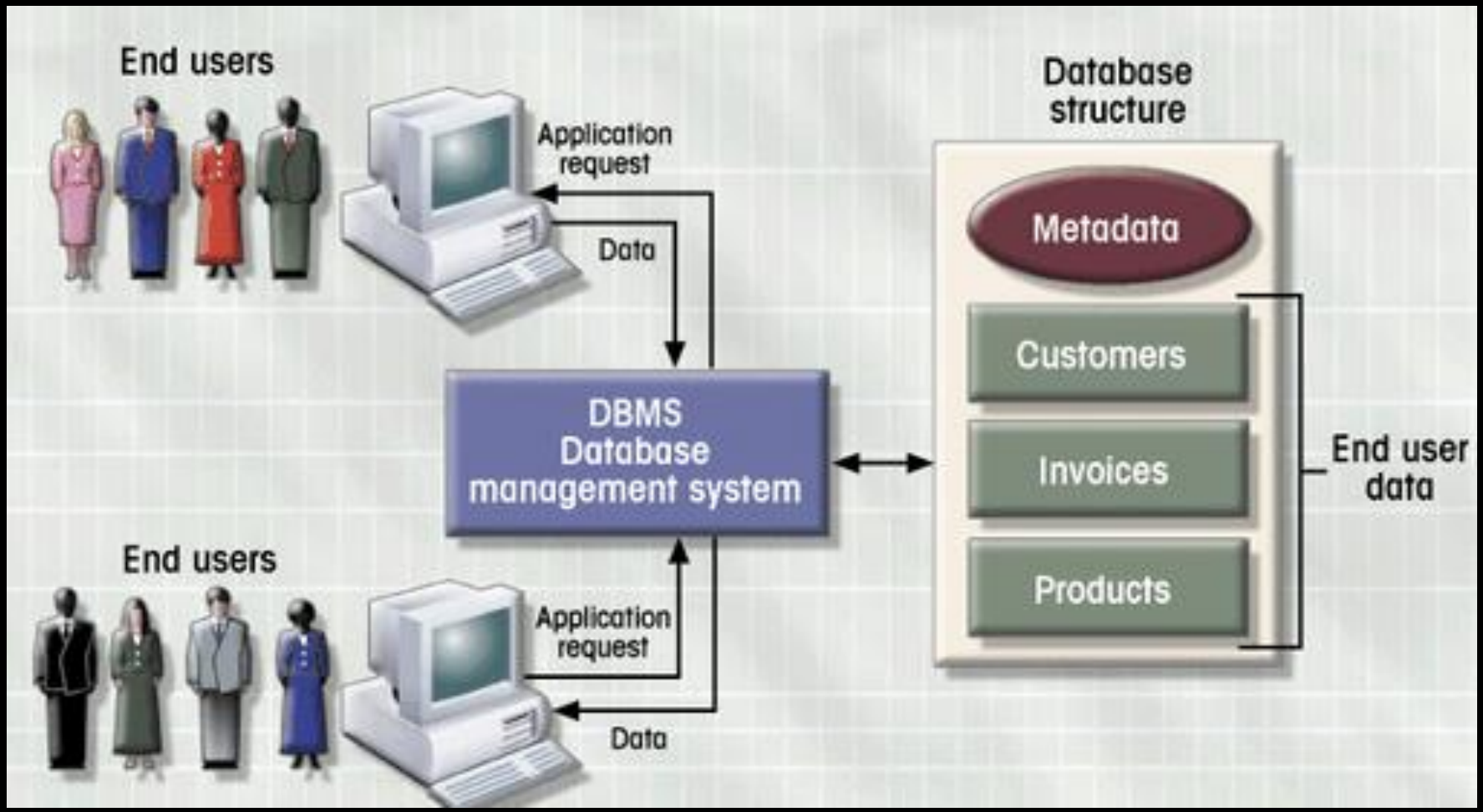
File Processing versus Database

File processing system	Database approach
• Each department has its own set of files • Used for many years • Have data redundancy • Isolate data	• Programs and users share data • Reduce data redundancy • Improve data integrity • Share data • Allows easier access • Reduces development time • Can be more vulnerable

Database vs file system



The DBMS


End users
Application request
Data
Database structure
Metadata
Customers
Invoices
Products
End user data
DBMS Database management system
End users
Application request
Data

The DBMS

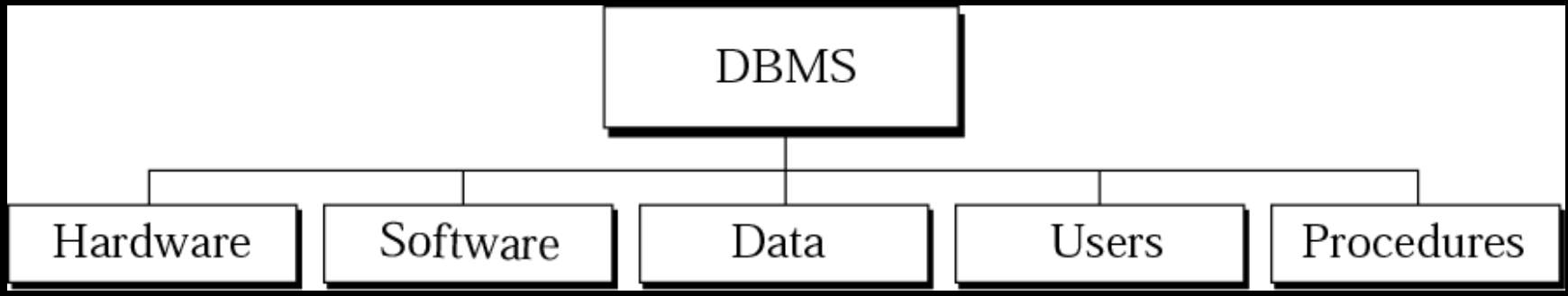
- The DBMS serves as the bridge between the user and the database. DBMS allows the user to manage, control or provide access to the data of the database.
- The DBMS manages the interaction between the end user and the database. DBMS presents single, integrated view of the data in the database.
- The DBMS receives all application requests and translates them into the complex operations to fulfil those requests. The DBMS hides much of the database's internal complexity from application programs and users.
- The application programs are usually written languages such as VB, Java, C++ and etc. DBMS enables the data in the database *to be shared* among multiple applications or users. DBMS *integrates* the many different users' views into a single data repository.

Query Examples

- SQL

```
Select FirstName, LastName, Phone
From tblPhoneNumbers
Where LastName=“Norton”;
```

Components of the DBMS Environment



Components of the DBMS Environment

- Hardware
  - The physical computer system that allows physical access to data
- Software
  - The actual program that allows users to access, maintain, and update physical data
- Data: stored physically on the storage devices
- User: include end users and application programs
- Procedure
  - A set of procedures (rules) that should be clearly defined and followed by the users of the database

Procedures

- Procedures refer to the instructions and rules that govern the design and use of the database.
- The users of the system and the staff that manage the database require documented
- procedures on how to use or run the system. These 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. This may include procedures on how to identify

The Data Administrator (DA)

- Responsible for the management of the data resource including database planning, development and maintenance of standards, policies and procedures, and conceptual/logical database design.
- The DA consults with and advises senior managers, ensuring that the direction of database development will ultimately support corporate objectives.

The **Database Administrator** (DBA)

- is responsible for the physical realization of the database, including physical database design(how computer sees the data) and implementation, security and integrity control, maintenance of the operational system, and ensuring satisfactory performance of the applications for users.
- The role of the DBA is more technically oriented than the role of the DA, requiring detailed knowledge of the target DBMS and the system environment.
- In some organizations there is no distinction between these two roles; in others, the importance of the corporate resources is reflected in the allocation of teams of staff dedicated to each of these roles.

Database analysts and administrators are responsible for managing and coordinating all database activities

Database Analyst (**DA**)

Decides on proper field placement, defines data relationship, and identifies users' access privileges

Database Administrator (**DBA**)

Creates and maintains the data dictionary, manages security, monitors performance, and checks backup and recovery procedures

Database Designer

- The **logical database designer** is concerned with identifying the data (that is, the entities and attributes), the relationships between the data, and the constraints on the data that is to be stored in the database.
- The logical database designer(how human being views data) must have a thorough and complete understanding of the organization's data and any constraints on this data (the constraints are sometimes called **business rules**).

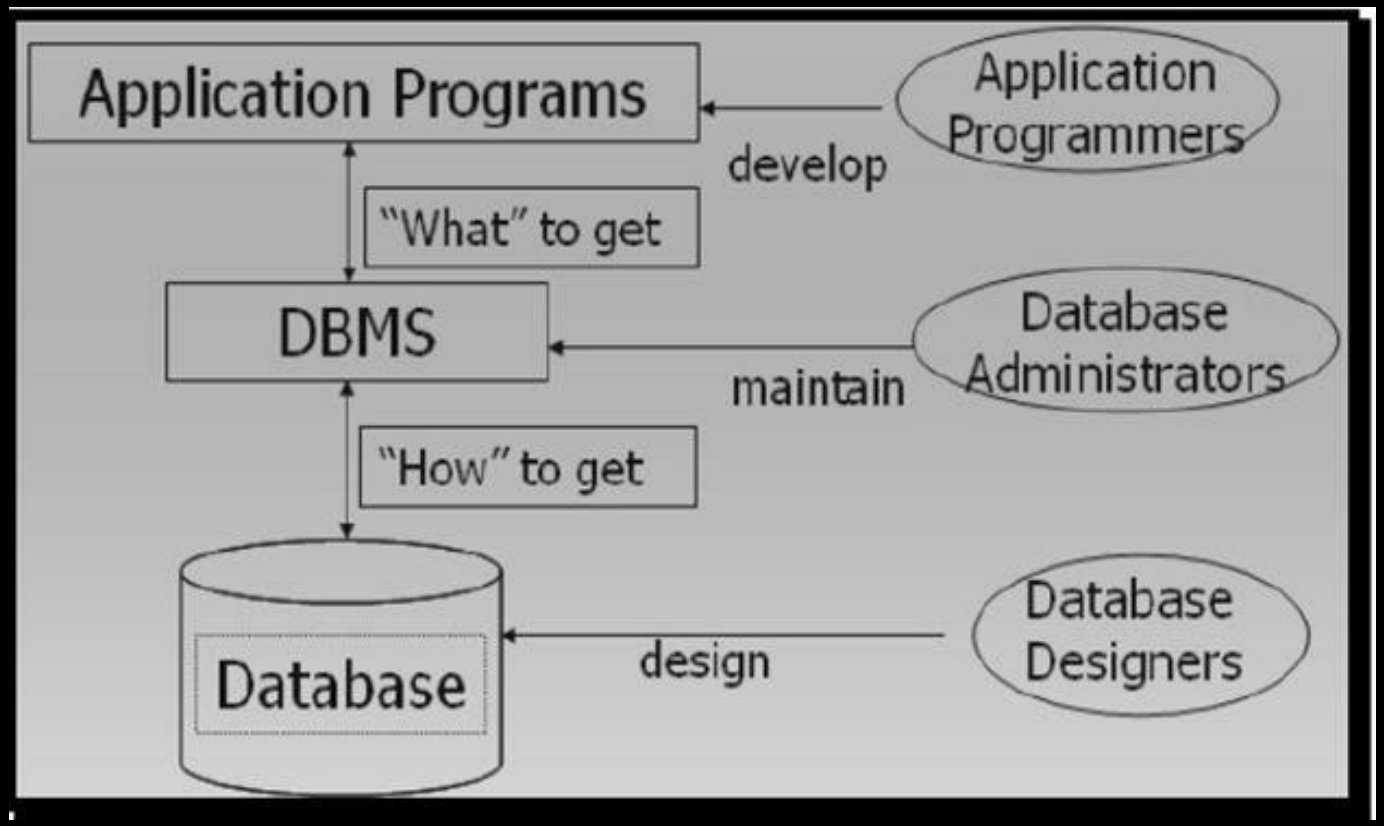
Physical database designer

- The **physical database designer** decides how the logical database design is to be physically realized.
- This involves:
  - mapping the logical database design into a set of tables and integrity constraints;
  - selecting specific storage structures and access methods for the data to achieve good performance;
  - designing any security measures required on the data.

Application Developers

- Once the database has been implemented
- The application programs that provide the required functionality for the end-users must be implemented. This is the responsibility of the **application developers**

Summary


Application Programs
Application Programmers
develop
"What" to get
DBMS
Database Administrators
maintain
"How" to get
Database
Database Designers
design

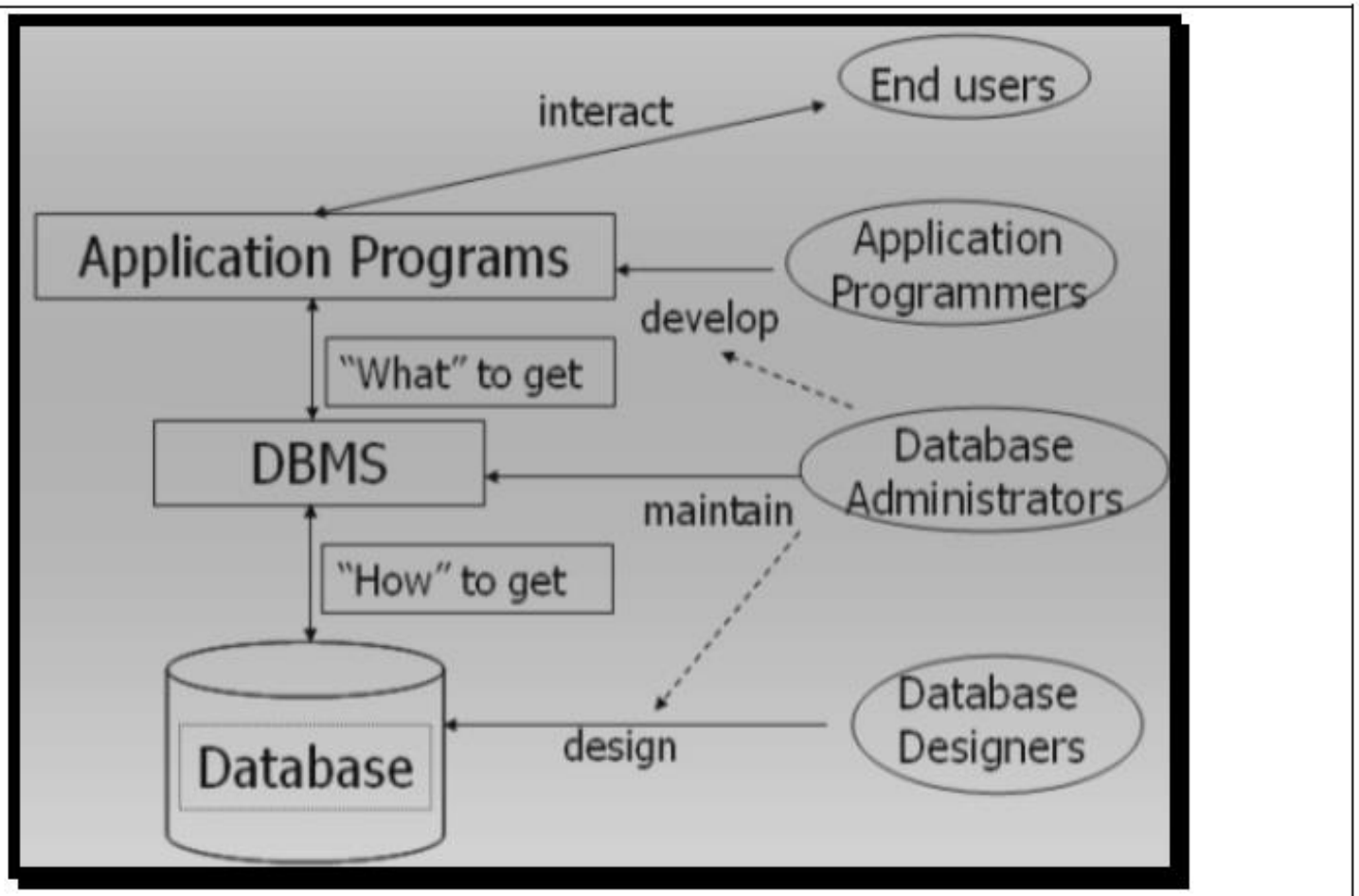


Fig. 7: Database Administration's interaction with other users

Advantage of DBMS

- Control of data redundancy
- Data consistency
- Sharing of data
- Improved data integrity
- Improved security
- Enforcement of standards
- Economy of scale
- Improved data accessibility and responsiveness
- Increased productivity
- Increased concurrency
- Improved backup and recovery services

Advantages of Database Approach

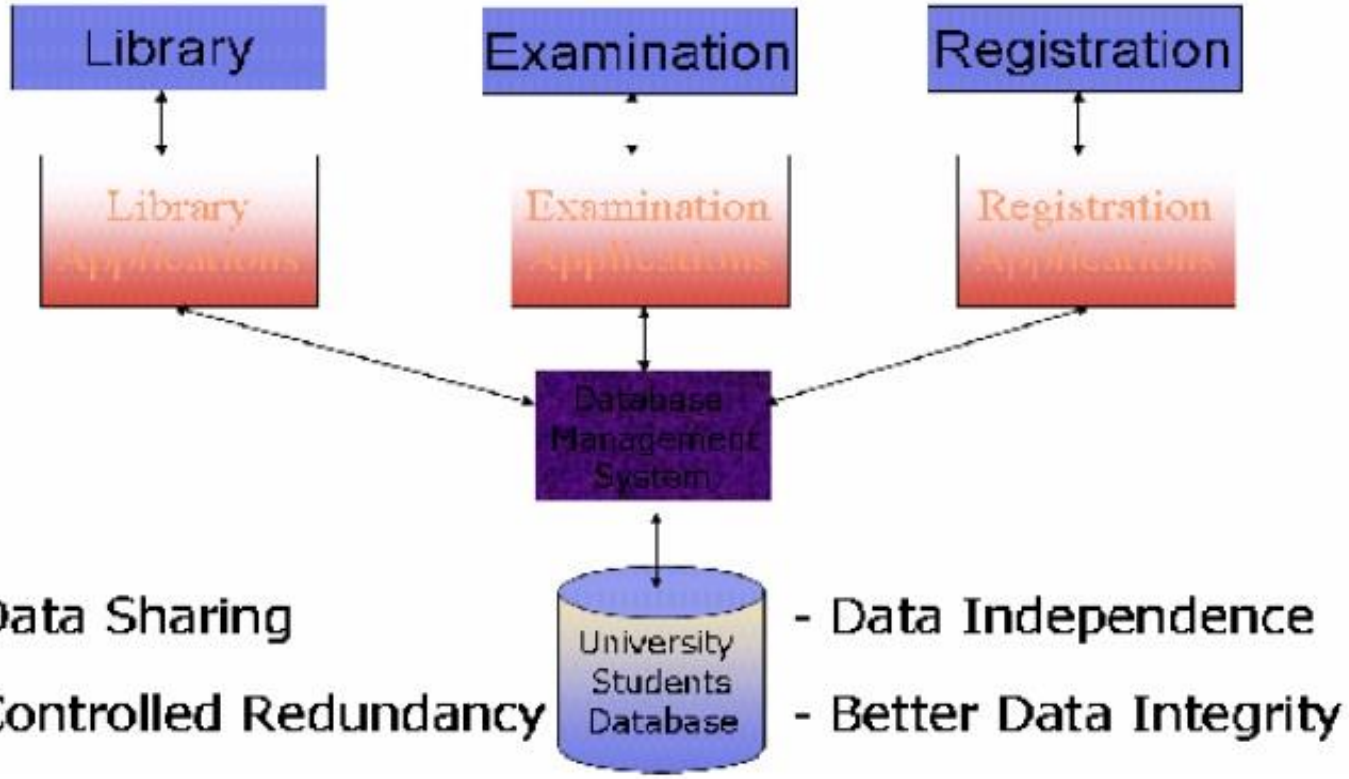


Fig. 3: A typical Database System environment

Disadvantage of DBMS

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

Common Corporate DBMS

- Oracle
  - Most popular enterprise-level DBMS
  - Very flexible storage system
  - Can be very complex
  - Platform independent
    - Running on various operating system
  - Offers a wide range of solutions
  - Handles trillon of bytes (tera bytes) simultaneously handle thousands of user

Common Corporate DBMS

- DB2
  - Venerable IBM database
  - Platform independent
  - Only database using pure SQL
- **IBM® DB2®** is the database of choice for robust, enterprise-wide solutions handling high-volume workloads.
- Provide security and scalability
- It is optimized to deliver industry-leading performance while lowering costs.
- Industry-leading performance for multi-work loads

Common Corporate DBMS

- Microsoft SQL Server
  - Fastest growing DBMS
  - Only runs on Microsoft platforms
  - Eight different versions exist
  - Extremely scalable architecture
    - Software can grow with the data

Common Corporate DBMS

- MySQL
  - Leading DBMS for Linux
  - inexpensive
  - Features are those needed in business
  - Often faster than other DBMS
  - Platform independent
  - It can be used for free
  - Popular for new internet database work