

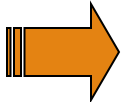
Chapter 1 Introduction

Contents

- The History of Database System
- Overview of a Database Management System (DBMS)
- Three aspects of database-system studies
- the state of the art

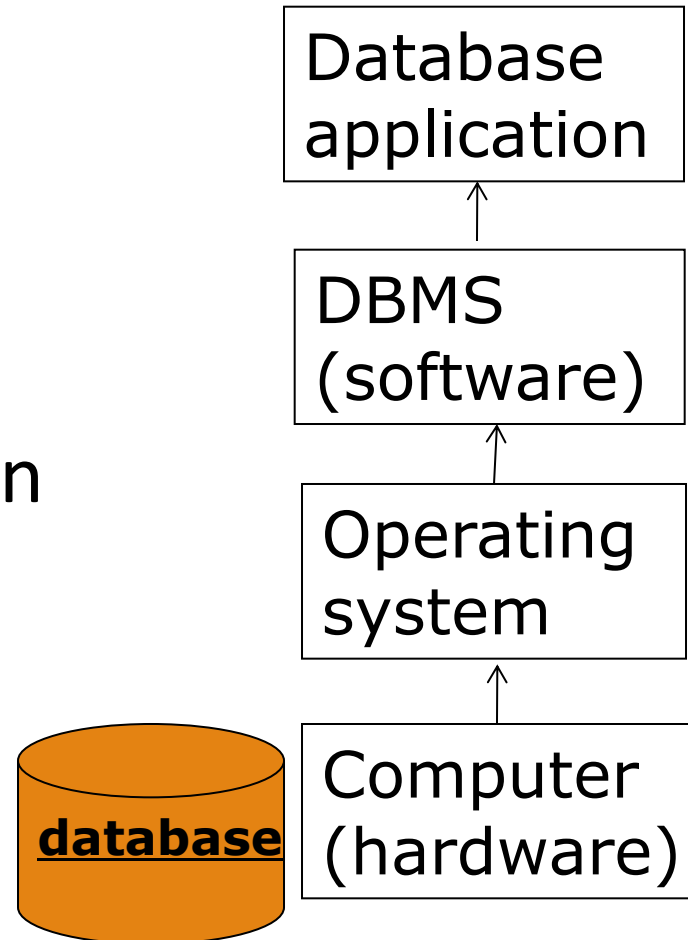
Introduction to Database Systems

What is a Database System?

- Database (data , metadata)
- Hardware (disks)
-  ■ Software (**DBMS**)
- People (users, database designers and database administrators DBA)

Introduction to Database Systems (cont.)

- Database Applications
- Database Management System (**DBMS**)
- Operating System (not in this course)



Database Application Systems

- **Databases touch all aspects of our lives**
 - Banking: all transactions
 - Airlines: reservations, schedules
 - Universities: COURSE registration, grades
 - Sales: customers, products, purchases
 - Manufacturing: production, inventory, orders, supply chain
 - Human resources: employee records, salaries, tax deductions ...

The history of DBMS

- 1950s and early 1960s:
 - Data processing using **magnetic tapes** for storage
 - Tapes provided only sequential access
 - Punched cards for input
- Late 1960s and 1970s:
 - Hard disks allowed direct access to data
 - **Network** and **hierarchical data models** in widespread use
 - Ted Codd defines the **relational data model**
 - he won the ACM Turing Award for this work
 - IBM Research begins System R prototype
 - UC Berkeley begins Ingres prototype
 - High-performance (for the era) transaction processing

More detailed History (cont.)

- 1980s:
 - Research relational prototypes evolve into commercial systems
 - **SQL becomes industrial standard**
 - Parallel and distributed database systems
 - **Object-oriented** database systems
- 1990s:
 - Large decision support and data-mining applications
 - Large multi-terabyte **data warehouses**
 - Emergence of Web commerce
- Early 2000s:
 - **XML and XQuery standards**
 - Automated database administration
- Later 2000s:
 - Giant data storage systems, noSQL systems
 - stream data, uncertain data,

The state of the art

from DTCC2014

大数据的宏观视图：行业与互联网大数据 DTCC2013

大数据领域

行业大数据

互联网大数据

经营类

电信信令
电信话单
金融细账
金融票据
电力调度
智能电网
经营分析

结构化为主

管理类

文件
报表
纳税分析
社保分析
决策支持
预测

结构化

+半结构化

监管类

公安网监
国安技侦
舆情监控
银监会稽查
食品溯源
环保监测

结构化

+半结构化

专业类

音视频
地震勘探
气象云图
卫星遥感
雷达数据
物联网

非结构为主

10%结构化

30%半结构化

60%非结构化

价值密度

结构化

>半结构化

>>非结构化

SQL

NewSQL

NoSQL

What is a DBMS (database management system) ?

From Commercial Market:

- Oracle, Sybase are among the largest software companies in the world.
- IBM offers its DB2 system
- Microsoft offers SQL-Server & Microsoft Access for DBMS on desktop
- MySQL

What do you think about DBMS?

From User's point of view:

Collection of interrelated data

Set of programs to access the data

DBMS contains information about a particular enterprise, *reliable* and *safe*.

DBMS provides an environment that is both *convenient* and *efficient* to use.

The DBMS is expected to:

- Allow users to **create** new database and specify their structure.
- Give users the ability to **query** the data.
- Support the **storage** of very large amounts of data.
- **Enable durability, recovery** of failures, errors of many kinds.
- **Control access** to data from many users at once.

DBMS (From its Functions)

Design the structure of their information (e.g. relational model)

Operate (query, modify) on the data

Manage huge amount of data and support **efficient, concurrent, secure, atomic access** to very large amounts of data

DBMS (from implementation, how to manage data)

Persistent storage: supports the storage of very large amounts of data:

1. **Data**: the content of DB itself.
2. **Metadata**: describes the structure of, and constraints on, the database.
3. **Log records**: recent change to the databases.
4. **Statistics**: about data property, such as the sizes of, values of,...
5. **Indexes**: support efficient access to the data

Programming interface: to access and modify data through a powerful query language.

DBMS: how to manage data (cont.)

Transaction management:

supports **secure, atomic access** to very large amounts of data.

Concurrent control:

supports **concurrent access** to data.

DBMS (three components)

Query processor:

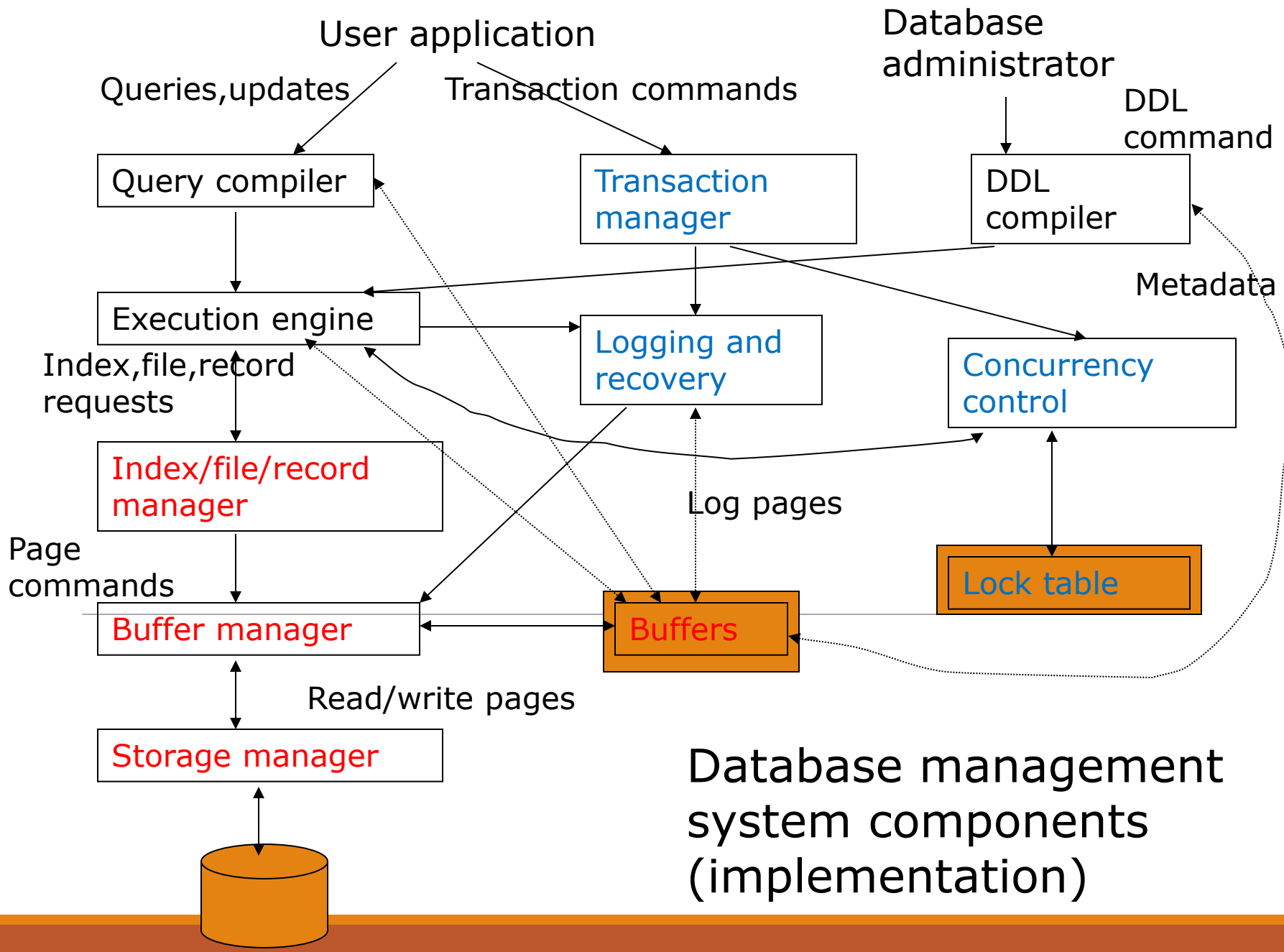
The query compiler, execution engine

Transaction manager:

Logging, concurrency control, deadlock resolution

Storage manager:

Control the placement of data on disk, and its movement between disk and main memory



Database management system components (implementation)

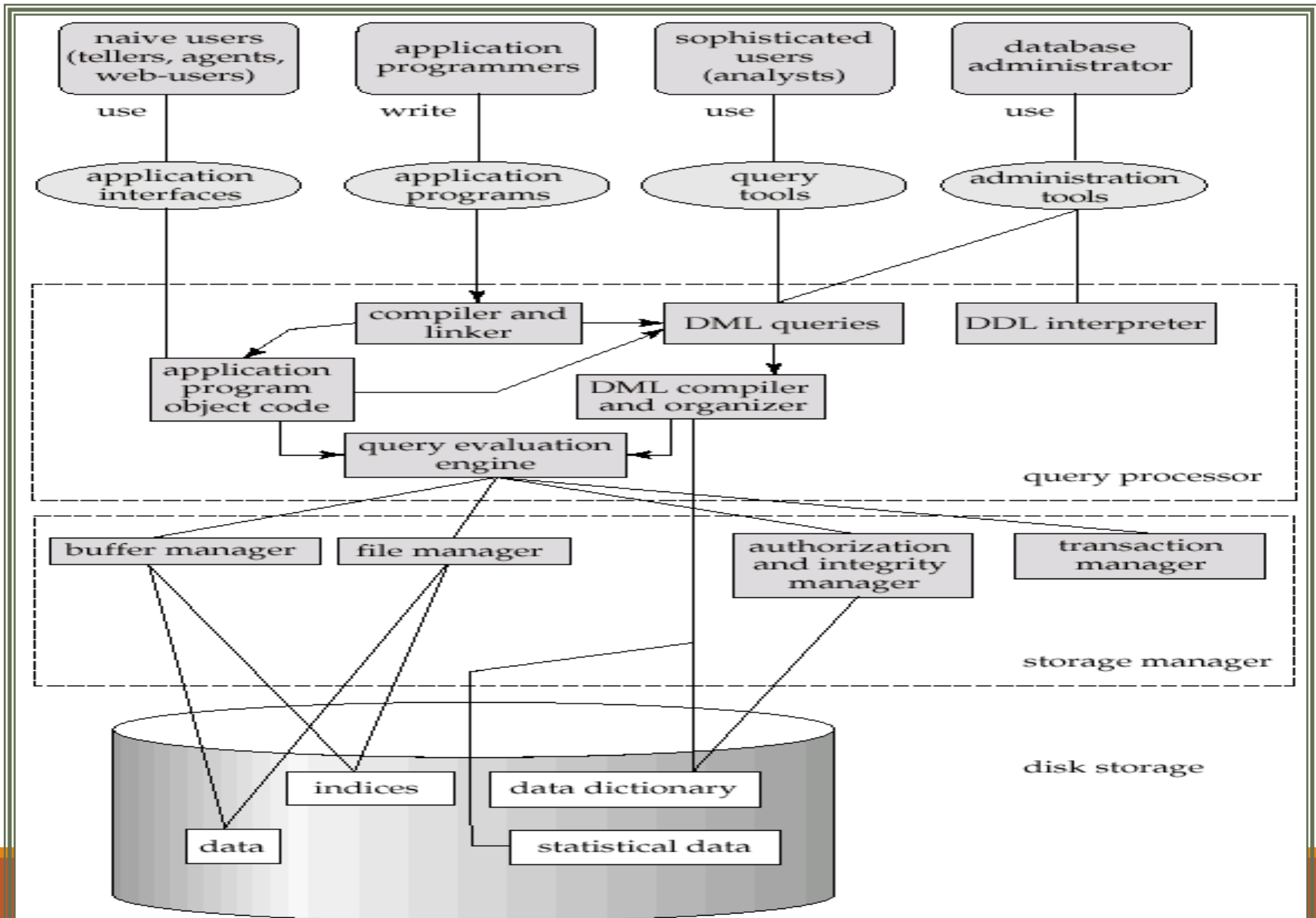
Many Users

- **Application programmers** – interact with system through database manipulation language
- **Sophisticated users** – form requests in a database query language
- **Specialized users** – write specialized database applications that do not fit into the traditional data processing framework
- **Naïve users** – invoke one of the permanent application programs that have been written previously
 - 👉 E.g. people accessing database over the web, bank tellers, clerical staff

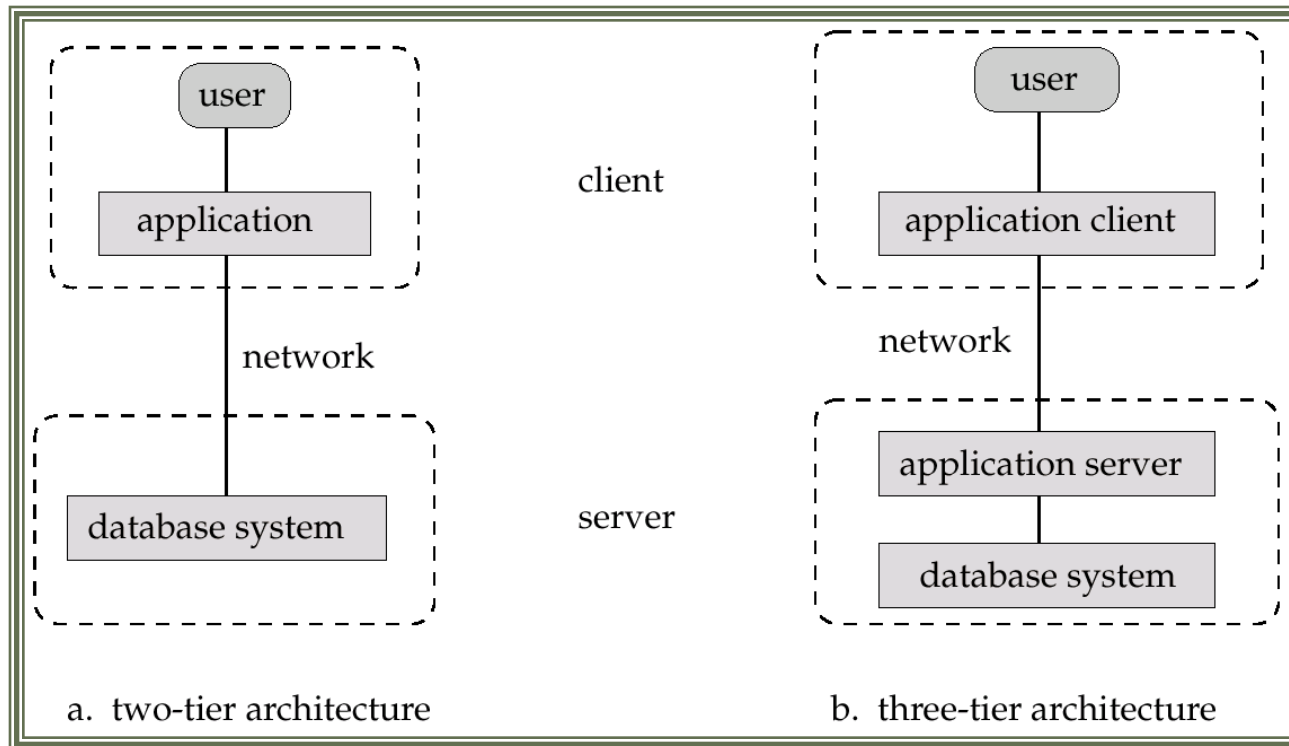
Database Administrator (DBA)

- Coordinates all the activities of the database system; the database administrator **has a good understanding** of the enterprise's information resources and needs.
- Database administrator's **duties** include:
 - Schema definition
 - Storage structure and access method definition
 - Schema and physical organization modification
 - Granting user authority to access the database
 - Specifying integrity constraints
 - Acting as liaison with users
 - Monitoring performance and responding to changes in requirements

Different Users with the System



Application Architectures



- **Two-tier architecture:** E.g. client programs using **ODBC/JDBC** to communicate with a database
- **Three-tier architecture:** E.g. web-based applications, and applications built using “middleware”
- **Multi-tier architecture.**

Three Aspects of Studying DBMS's

Modeling and design of databases. (**how is the information structured?**)

Programming: queries and DB operations like update. (**how does one express queries and other operations on the database**)

DBMS implementation. (how to build a DBMS)

Summary of chapter 1

- Database System & Applications
- Database Management System (DBMS)
features, functions, components (storage manager, the query processor, and the transaction manager)
- Database Language (DDL, DML)
- The development of DBMS

Past :

Network DBMS

Hierarchical DBMS

Relational DBMS

Now:

Relational DBMS

Object-Relational/oo DBMS

Web-based database,
noSQL system

Homework

Read section 1.2