

IMS processes the world's transactions



50,000,000,000
secure transactions per day

Introduction to Databases

Database

- Database is collection of related data
 - Data is collection of known facts with explicit meaning
- Database Management System (DBMS)
 - Collection of programs that enable users to create, maintain, and access a database

Why do we need databases?

- ❑ We don't need databases
- ❑ Data could be stored in files
- ❑ Applications could use custom code that accesses/modifies these files
- ❑ Example: an application that keeps track of student data, like class schedules

Problems with files and custom code

- ❑ You want a new feature to show all students who took a class in a specific semester – now you need to write code to find that data
- ❑ How to avoid data loss if a file becomes corrupted or your code has a bug?
- ❑ How to support lots of users on the system at the same time?
- ❑ How to support permissions for students, faculty, administrators?
- ❑ How to provide good performance as data grows?

Solving the problems

It is possible to solve all of these problems with custom code.

However:

- they are difficult problems
- lots of applications need to solve the same problems

Database systems

A database system is a general-purpose data management system that supports most or all of these features:

high
performance

access control
(permissions)

general-
purpose data
querying
language

reliability
99.99% up
time

transaction
support

support for
concurrent
data access

general-
purpose data
definition
language

protection
against data
loss and
corruption

physical data
independence

handle data at
massive scale

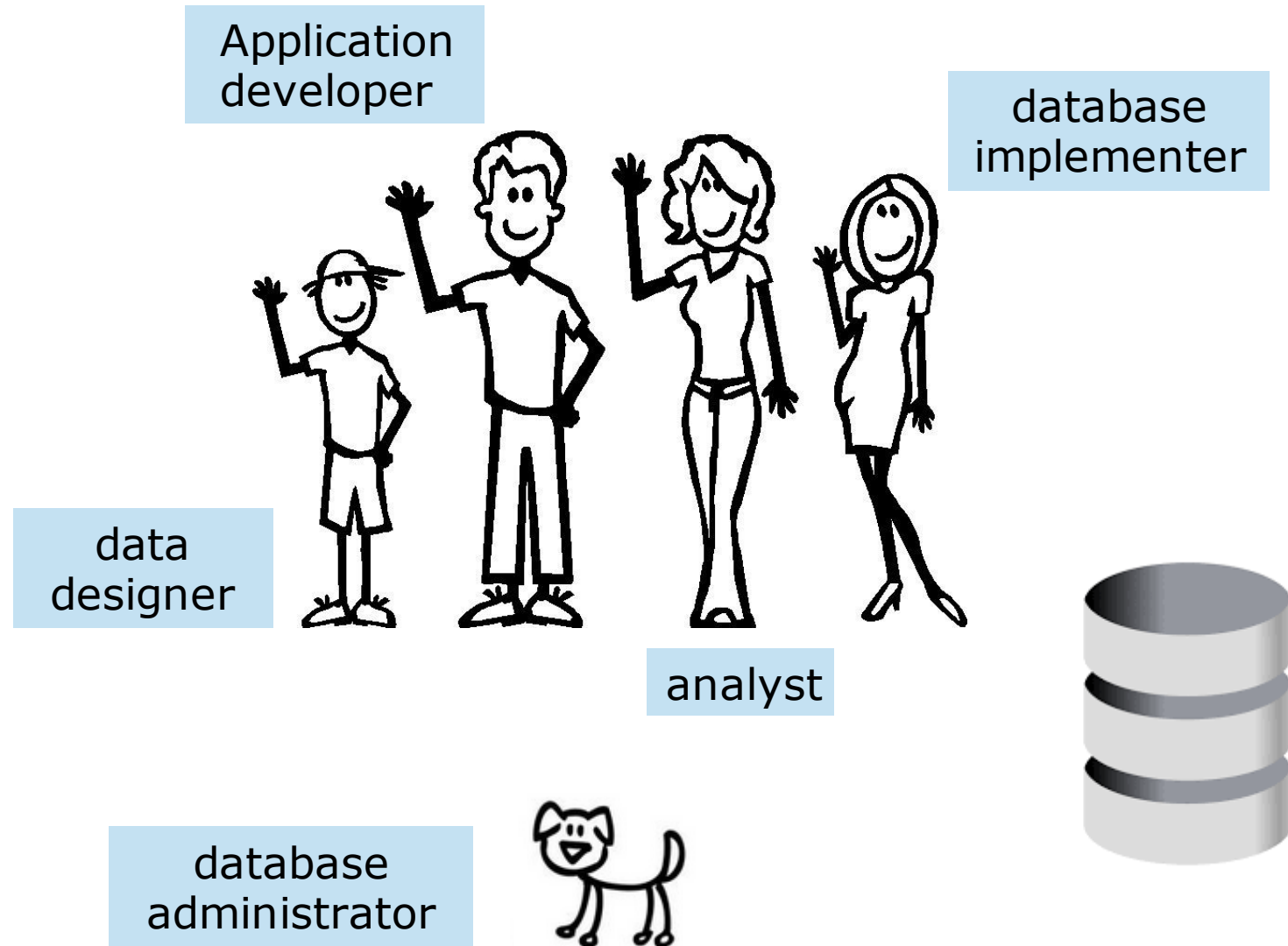
When not to use a database

Databases are complex systems and incur lots of overhead.

Consider not using a database system when:

- ☐ the data and applications are simple and well-defined
- ☐ there are tight real-time requirements
- ☐ multiple user access is not required

Roles around databases



Database concept: data model

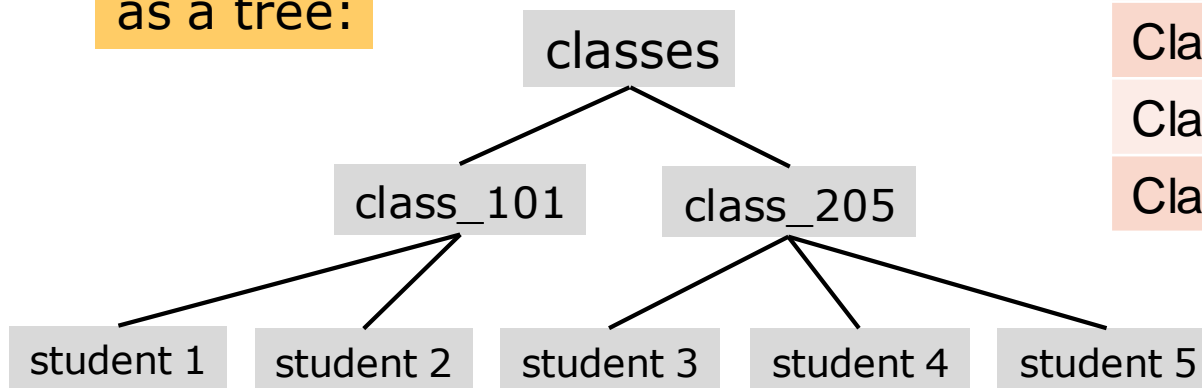
A **data model** is a concept for data organization (how the data is structured).

For example, we could store student data as a tree, a list, a graph, or a table.

as a table:

class	student
Class_101	Student_1
Class_101	Student_2
Class_205	Student_3
Class_205	Student_4
Class_205	Student_5

as a tree:



Database schema

A **schema** is the design or structure of a specific database.

An **instance** is a schema “instantiated” with data

schema:

```
table roster {  
  class: string,  
  student: string  
}
```

instance:

class	student
Class_101	Student_1
Class_101	Student_2
Class_205	Student_3
Class_205	Student_4
Class_205	Student_5

Database queries

A **query** is a statement requesting the retrieval of information from a database.

Declarative: describe what you want out of the database without describing the algorithm to get the data out.

Example SQL queries:

```
SELECT * FROM Student WHERE Country='Norway';
```

```
SELECT * FROM Student WHERE City='San Jose' OR City='Seattle';
```

Transactions

A **transaction** is the unit of change in a database.

Transactions contain a set of database operations.

A transaction must succeed, or fail “completely”.

For example, a transaction might be used for moving a student from one class to another.

Lecture Lab

1. In your own words, and without looking at the lecture notes, define “database”.
2. A database is a general-purpose piece of software that is used within many kinds of applications. Can you name one such general-purpose pieces of software?
3. Name and define four important features of a database.
4. Explain the job of a database administrator.
5. Is data stored in a schema or an instance?

Lecture Lab - solutions

1. First, it is a system for data storage. Second, it is general purpose in that it includes a general-purpose way to describe data, and a general-purpose way to query data.
2. Operating system
3. Concurrent access, access control, general-purpose query language, high performance, high reliability, support for transactions.
4. The job typically things like database design, database migration, setting up and managing access control, database maintenance (like backup), database tuning, and managing the physical infrastructure.
5. Data is stored in an instance. A schema describes how data is structured. You can think of a schema as the 'type' of a table -- it give the attribute names and types.