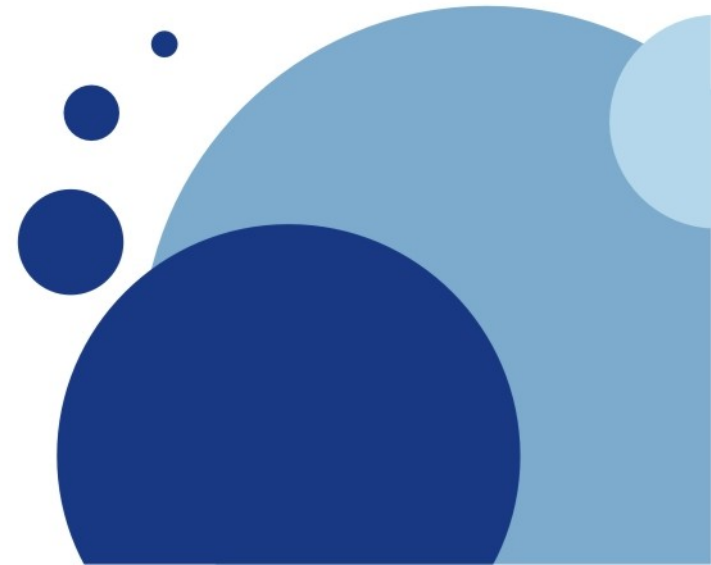# Introduction to the Worlds of Database Management Systems

Dr. Rim Moussa
University of Carthage

Lecture Outline

- Introduction
- Entity-Relationship Diagram
- Logical Relational schema
- Relational Algebra
- Physical Relational Schema
- RDBMS Architecture
- Structured Query Language
- Query Optimization
- RDBMS Projects 2020
  - Airline Cie database
  - CRM for ENI-Carthage database
  - BTC Transactions database
  - NYC mobility and PoI database
  - Retail broker database

- Databases are essential to every business
- What is a database?
  - A database is a collection of information that exists over a long period of time, often many years. In common parlance, the term database refers to a collection of data that is managed by a DBMS. (H. Garcia-Molina, J. Ullman, J. Widom)
- Corporations maintain all their important records in databases
  - e.g. Bank, Airline Cie, Library, Hospital/Medical center, School/ University, Hotel, Insurance Cie, Retail store, Civil registrar, telecommunication, navigation systems, census db…
- Different types of databases for different usages
  - graph database: Social DB
  - document database: Text DB
  - time-series database: IoT data: weather
  - spatial database: Spatial data
  - Structured data: relational database

- An example of a large commercial database is *Amazon.com*. It contains data for over 60M active users, and millions of items.
  - The database occupies over 42TB and is stored on hundreds of computers (called servers). Millions of visitors access *Amazon.com* each day and use the database to make purchases.
  - The database is continually updated as new books and other items are added to the inventory, and stock quantities are updated as purchases are transacted.
  - Amazon RDS
- There are two modes in which databases are used.
  - *On-line transaction processing* -where a large number of users use the database, with each user retrieving relatively small amounts of data, and performing small updates.
  - *On-line analytical processing* -the processing of data to draw conclusions, and infer rules or decision procedures, which are then used to drive business decisions.

DBMS: Database Management System

- A DBMS is a computerized system that enables users to create and maintain a DB.

- The DBMS is a general-purpose software system that facilitates the processes of defining, constructing, manipulating, and sharing DBs among various users and applications. (R. Elmasri and S. B. Navathe)

- The primary goal of a DBMS is to provide a way to store and retrieve database information that is both *convenient* and *efficient*. (A. Silberschatz, H. F. Korth and S. Sudarshan)

DBMS: Database Management System

- (H. Garcia-Molina, J. Ullman, J. Widom) The DBMS is expected to:
  - Allow users to create new databases and specify their schemas using a specialized data definition language.
  - Give users the ability to query the data and modify the data, using an appropriate language, often called a query language or data manipulation language.
  - Support the storage of large amounts of data over a long period of time, allowing efficient access to the data for queries and DB modifications.
  - Enable *durability*, the recovery of the database in the face of failures, errors of many kinds, or intentional misuse.
  - Control access to data from many users at once via a data control language and concurrency control for transactions,

Relational Database Management Systems

- Model
  - Instance > database > table > row
- Query Patterns
  - Selection based on complex conditions, projection, equi-joins, theta-joins, aggregation, derivation of new values (calculated attributes), recursive queries, nested queries, …
- Query languages
  - SQL (Structured Query Language)
    - Standard of the American National Standards Institute (ANSI) in 1986, and of the International Organization for Standardization (ISO) in 1987.
  - Formal: Relational algebra, relational calculi (domain, tuple)
- Implementations

- Database applications are typically broken up into a front-end part that runs at client machines and a part that runs at the backend. In two-tier architectures, the front end directly communicates with a database running at the back end. In three-tier architectures, the back end part is itself broken up into an application server and a database server.

JAVA: JDBC
MS Visual C++:
ODBC, ADO.NET

(a) Two-tier architecture

(b) Three-tier architecture

8
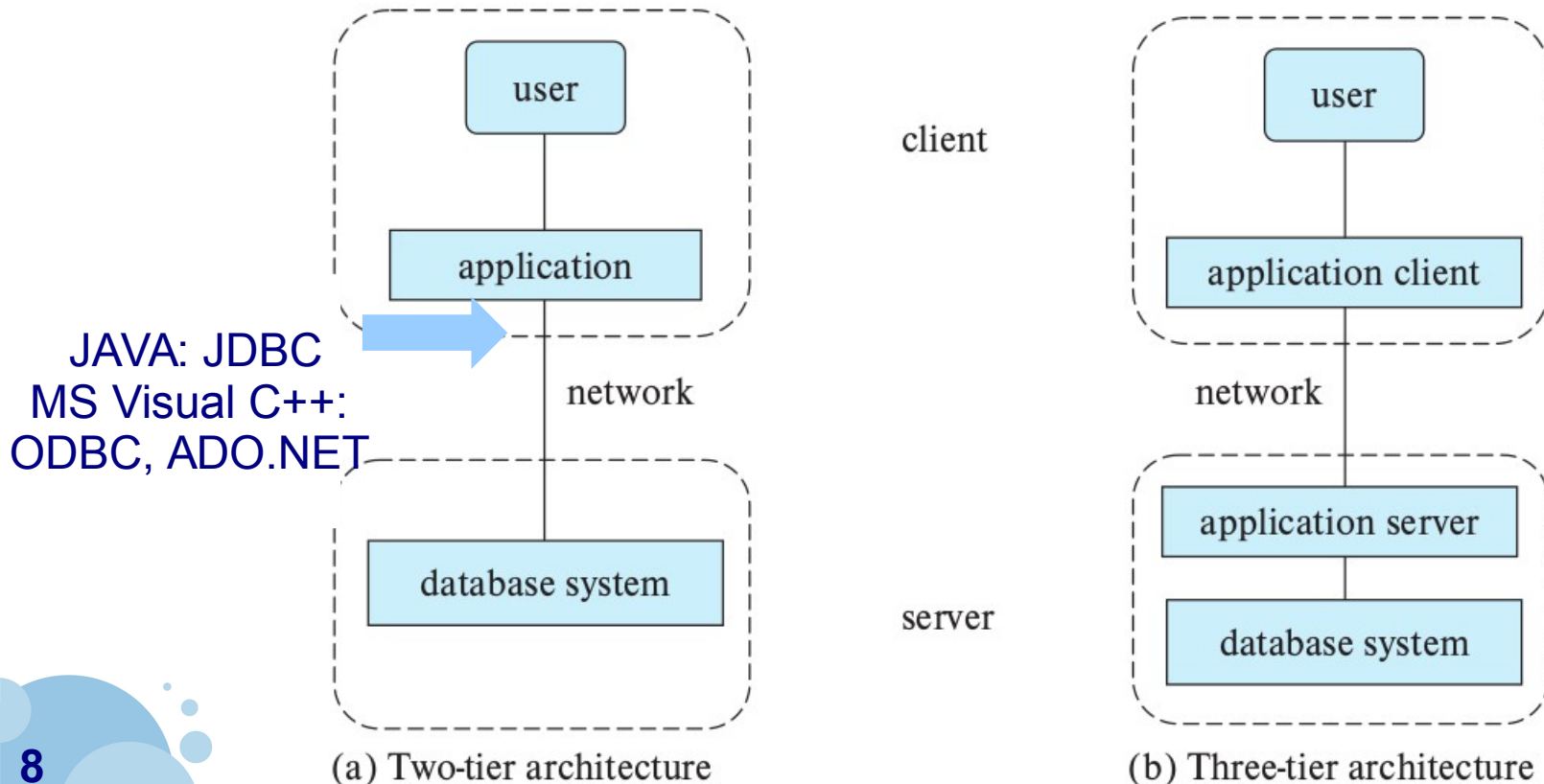
References

- A. Silberschatz, H. F. Korth and S. Sudarshan: Database System Concepts, Mc Graw Hill Education, 7th Edition 2020.

- H. Garcia-Molina, J. Ullman and J. Widom: Database Systems -the complete book, 3rd edition, Prentice Hall 2008

- R. Elmasri and S. B. Navathe, Fundamentals of Database Systems, 7th edition, Addison Wesley 2015.