Database Programming with JDBC
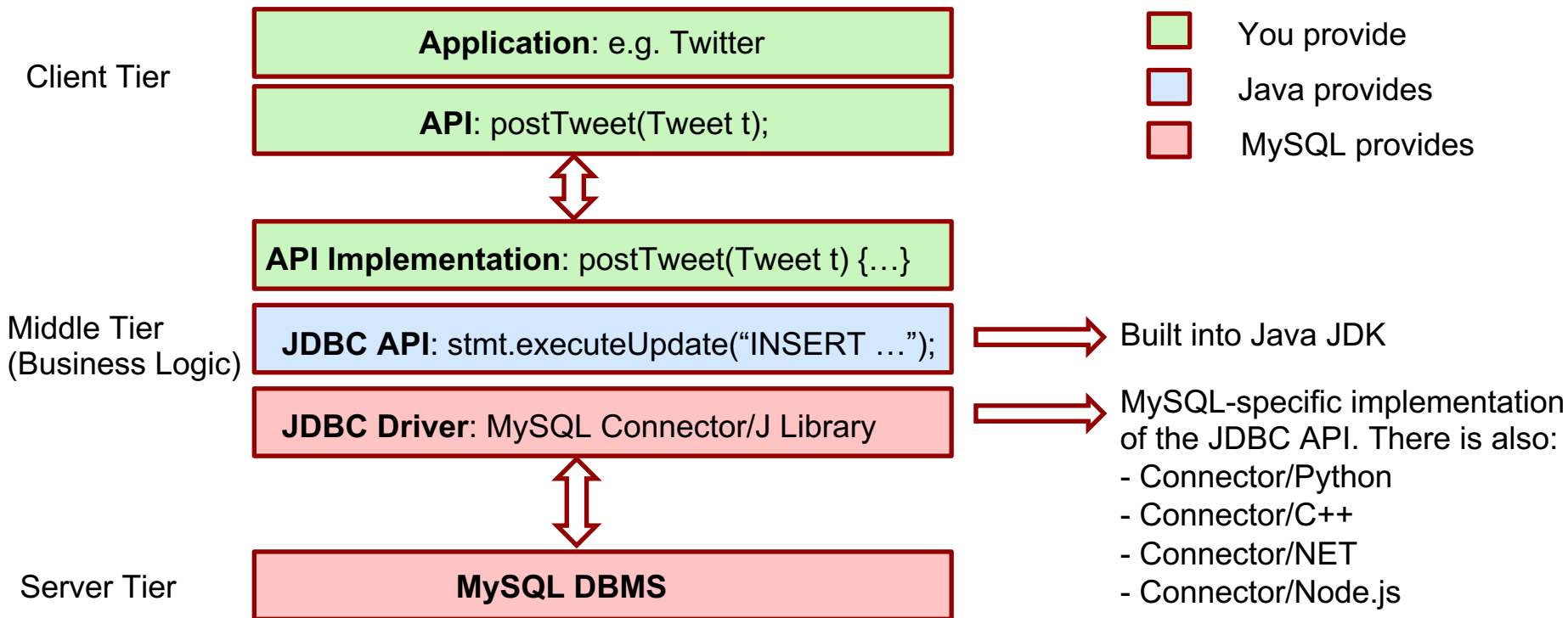
CS3200: Database Design
CS5200: Database Systems

Database Application
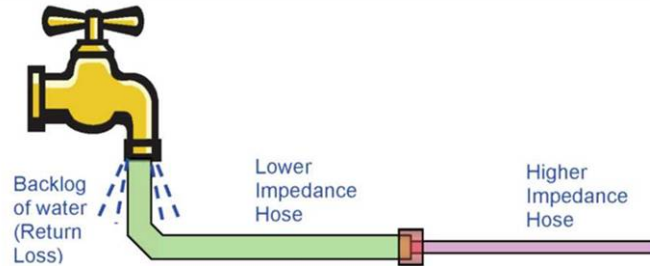
APIs and Drivers

Client Tier

| Application: e.g. Twitter |
| API: postTweet(Tweet t); |

Middle Tier (Business Logic)

| API Implementation: postTweet(Tweet t) {...} |
| JDBC API: stmt.executeUpdate("INSERT ..."); |
| JDBC Driver: MySQL Connector/J Library |

Server Tier

| MySQL DBMS |

■ You provide
■ Java provides
■ MySQL provides

→ Built into Java JDK

→ MySQL-specific implementation of the JDBC API. There is also:
- Connector/Python
- Connector/C++
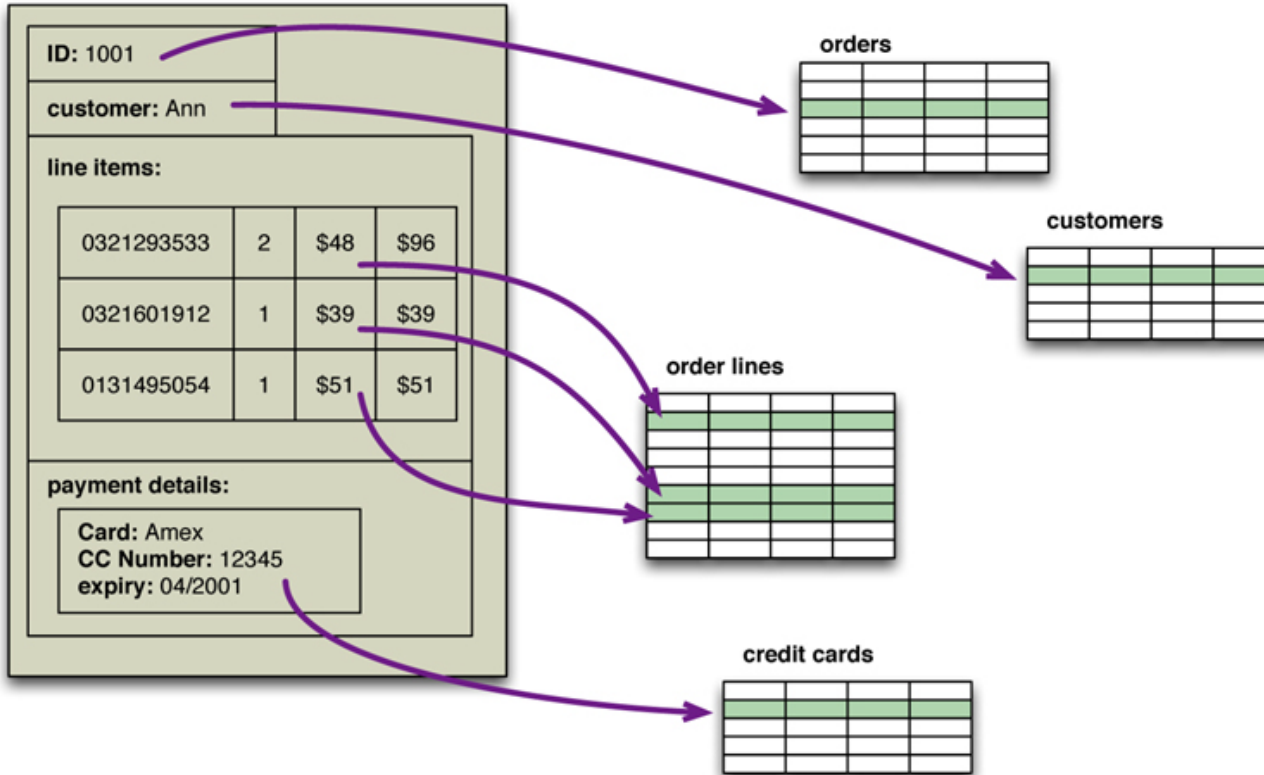- Connector/NET
- Connector/Node.js
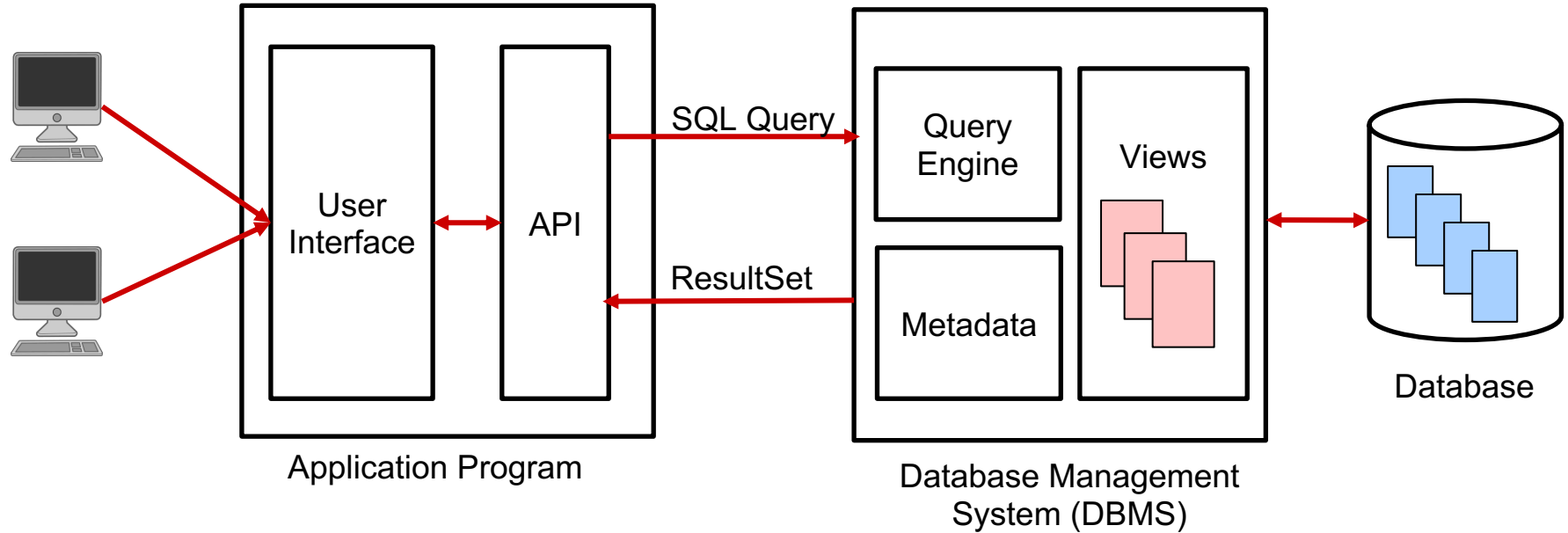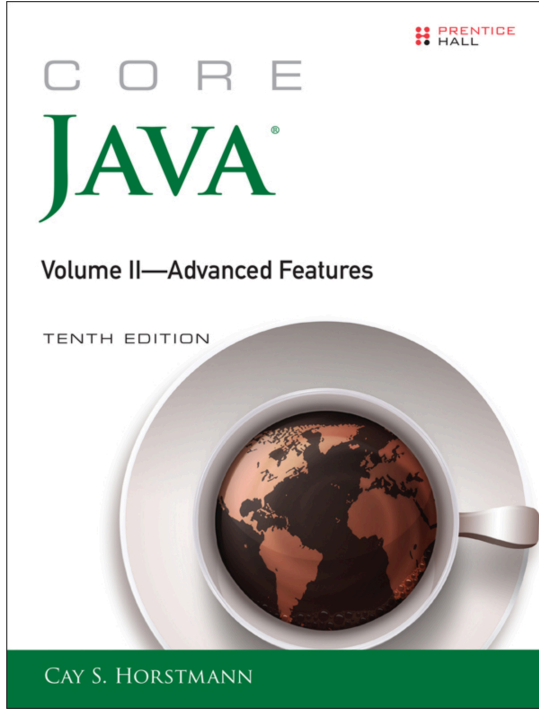
Impedance Mismatch

In this context, refers to several issues that arise when OO language interacts with RDBMS

- ○ Differences in data types
- ○ Query results as row/column
- ○ Limited compile-time error detection w.r.t. SQL

Impedance mismatch problem

Recommended Reading

- **Chapter 5**. Database Programming (Sections 1-5, 8)

- Freely available through Safari / ProQuest

Typical Programming Sequence

1. Connect to DBMS
   - URL, database name, user/pw, driver
   - Sometimes *persistent* for performance

2. Arbitrary interactions
   - Transactions via SQL / JDBC calls

3. Close the connection

Query Sequence

1. Generate SQL
   - Could be static or composed of algorithmic/user-contributed parts

2. Execute: ResultSet rs = stmt.execute(sql)

3. Process results
   while (rs.next() != null) { … }

Prepared Query Sequence

1. Generate parameterized SQL
   ○Could be static or composed of algorithmic parts (typically nothing user-contributed)

2. Bind values to SQL parameters
   ○Could be static or algorithmic/user-contributed

3. Execute

4. Get results

Stored Programs vs. API-driven Application

	Stored Program	API-Driven Application
Technology Stack	MySQL	MySQL, JDBC, Java
Business Logic	Built-in	Separated out
Schema Change Impact	Store-programs might change but API calls to stored programs unaffected.	API Implementation might need to change, but API *interface* should be unaffected
Developer Talent	Database developers	Database developers Software developers
Impedance mismatch	Avoided. Everything done inside the DBMS	Must be tackled by the software developers
Vendor lock-in	High: Stored programming dialects vary - may be difficult to migrate to a new vendor	Low: JDBC relies more on SQL statements that are more standard.

Data Science vs. Data Engineering

Data Engineering   ## Data Science

Data Scientist: Performs descriptive statistics to develop insights, build models, solve a business need.

Data Engineer: Architect large scale data processing systems typically using relational, NoSQL, messaging, and distributed computing platforms.

Develop, construct, test, and maintain architectures (such as databases and large-scale processing systems)

Ensure architecture will support the requirements of the business

Discover opportunities for **data acquisition**

Develop data set processes for data modeling, mining and production

Employ a variety of languages and tools (e.g. scripting languages) to **marry systems together**

Recommend ways to **improve data** reliability, efficiency and quality

Conduct research to **answer industry and business questions**

Leverage large volumes of data from internal and external sources to answer that business

Employ sophisticated analytics programs, machine learning and statistical methods to **prepare data for use in predictive and prescriptive modeling**

Explore and examine data to **find hidden patterns**

Automate work through the use of predictive and prescriptive analytics

Tell stories to key stakeholders based on their analysis

Diverse technologies with blurred boundaries