

## Exercise 8

Released: June 12, 2018 · Discussion: June 18, 2018

### 1 Parallelizing Queries

Which kinds of parallelism do you know and how can they be used to speed up query processing?  
Given are the three tables Customer, Product and Sales:

- Customer(cid INT PRIMARY KEY, Name VARCHAR(40), Firstname VARCHAR(40), Age INT, Gender CHAR);
- Sales(sid INT PRIMARY KEY, cid INT FOREIGN KEY, pid INT FOREIGN KEY, Count INT, RetailPrice FLOAT);
- Product(pid INT PRIMARY KEY, Description VARCHAR(60), Category VARCHAR(40), Price FLOAT);

Discuss how the following queries can be parallelized. For which techniques, presented in the lecture, do you expect a performance gain and why?

1. `SELECT * FROM Customer WHERE Customer.Age > 30 AND Customer.Gender='M';`
2. `SELECT SUM(RetailPrice) FROM Sales;`
3. `SELECT Name, Firstname FROM Customer NATURAL JOIN Sales NATURAL JOIN Product WHERE Customer.Age < 40 AND Product.Category='Television';`

### 2 Concurrency Control

Given the following schedule:

$$S = \langle r_3(x), r_1(x), r_2(x), w_1(y), r_2(x), w_3(x) \rangle .$$

Draw the **serialization graph** for  $S$ . Is  $S$  **conflict serializable**? Give an explanation for your answer.