Dept. of CSE, BUET
CSE453, Jan 2021
CT2, Full marks: 20
Duration 25 minutes + 5 minutes for upload

Answer the following questions

Q1.: 8+2

Following relations are given:

Customer (c-id, name, age, income), with 0.5 million tuples
Sale (c-id, item-id, date, quantity, unit-price, total-price), with 2 million tuples

There are five nodes (N1, N2, N3, N4 and N5) in the parallel database system. The partition vector is as follows:
c-id = [111111, 222222, 333333, 444444]

Explain showing the detailed steps how the following queries will be executed with appropriate number of nodes.

a. SELECT * FROM customer as c, sale as s WHERE c.c-id = s.c-id
b. SELECT * FROM customer as c WHERE c.c-id <= 300000

Q2.

For the schema of question Q1, how will the following SQL be executed using pipeline parallelism by nodes N3, N4 and N5? 5

SQL: SELECT * FROM customer as c, sale as s WHERE c.c-id = s.c-id AND c.c-id >= 333333

Q3.

The SQL on the schema of Q1. is given as follows: 5

SELECT * FROM sale ORDER BY quantity

The partition vector quantity = [10, 30, 60, 100]
Explain with figure, the steps of processing the above SQL using exchange operator.