University of Rajshahi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

B.Sc. Engineering Part IV Even Semester Exam-2024

C. Code: CSE4262 & CSE4262 Title: Distributed Database Management System Lab

Time: 03 Hours **Full Marks: 17.5**

Question 1. (a) Create a database named **BankDB**. Under **BankDB**, create the following three tables with partitions for **Account** and **Transaction** tables partitioning on **branch_id**:

i. **Customer Table** (not partitioned)

- customer_id (Primary Key, BIGINT)
- name (VARCHAR(100), NOT NULL)
- email (VARCHAR(100), UNIQUE)

ii. **Account Table** (partitioned by branch_id)

- account_id (BIGINT, NOT NULL)
- customer_id (BIGINT, NOT NULL, FK → Customer)
- branch_id (INT, NOT NULL)
- balance (DECIMAL(12,2), default 0.00)
- Primary Key (account_id, branch_id)

iii. **Transaction Table** (partitioned by branch_id)

- transaction_id (BIGINT, Primary Key)
- account_id (BIGINT, NOT NULL, FK → Account)
- branch_id (INT, NOT NULL)
- amount (DECIMAL(12,2), NOT NULL)

(b) Insert **Customer.csv**, **Account.csv**, and **Transaction.csv** data into respective tables. If necessary you should use tmp table and delete the temp table after data insertion.

- **Customer:** 200 records → Customer.csv
- **Account:** 200 records → Account.csv
- **Transaction:** 500 records → Transaction.csv

(c) Write Hive commands to **add** a partition (branch_id=6) to the **Account and Transaction** tables.

(d) Write Hive commands to **drop** the partition (branch_id=3) from the **Account and Transaction** tables.

(e) Show the number of transactions in each partition.

Question 2: Find how many words of each length appear in a txt file.

Mapper: For each word, emit (word_length, 1).
Reducer: Sum up counts to get (word_length, total_count).

Example Input: *MapReduce makes big data processing easy.*

Mapper Output:	Reducer Output:
(9, 1) (5, 1) (3, 1) (4, 1) (10, 1) (4, 1)	(3, 1) (4, 2) (5, 1) (9, 1) (10, 1)

Question 3: Viva Voce.