

CSE 444 Practice Problems

Parallel DBMSs and MapReduce

1. Parallel Data Processing Algorithms

- (a) Describe how to compute the equi-join of two relations R and S in parallel. Describe how to compute a group by aggregation operation in parallel.

2. System Comparison

(a) In a parallel DBMS, why is it difficult to achieve linear speedup and linear scaleup?

(b) List two features common to a traditional DBMS and MapReduce.

Clarification: Here, “traditional DBMS” means a traditional *parallel DBMS*.

(c) List two features that are different between the two types of systems: i.e., features that are present in one system but not in the other. For example, you can give one feature present in MapReduce but absent in a parallel DBMS and one feature present in a parallel DBMS but missing from MapReduce (or any other combination).