Question 1. (0.50 points)

(a) What are the main software modules of a DDBMS (Distributed Database System)? Discuss the main function of each of these modules in the context of the client/server architecture.

(b) What are the main reasons for and potential advantages of distributed databases?

Question 2. (1.00 point)

Compare relational databases and the MapReduce/Hadoop/BigData Systems. Use the following criteria in your discussion:

(a) Structured and unstructured data support
(b) Cost of deployment
(c) Fault-tolerance
(d) Query expressiveness
(e) Transaction processing support

Question 3. (0.50 points)

Consider the following four schedules for transactions T1, T2 and T3:

a. r1(X); r3(X); w1(X); r2(X); w3(X);

b. r1(X); r3(X); w3(X); w1(X); r2(X);

c. r3(X); r2(X); w3(X); r1(X); w1(X);

d. r3(X); r2(X); r1(X); w3(X); w1(X);

(a) For each schedule, specify whether or not the schedule is conflict serializable and why.
(b) Then, for each serializable schedule, give the equivalent serial schedule.
Question 4. (0.50 points)
Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads objects X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y. Give two example histories over transactions T1 and T2 that result in a read-write conflict and a write-write conflict respectively. Note: two examples for each type of conflict (i.e., four examples in total).

Question 5. (0.50 points)
(a) Discuss the different types of transaction failures.
(b) What is meant by catastrophic failure?
(c) What is the system log used for?
(d) What are checkpoints and why are they important?