Databases and Distributed Computing Ph.D. Exam, sample questions

Answer any two questions. Partial credit will be awarded. This is a closed book one hour exam. Good luck.

Question 1

Can three phase commit produce an inconsistent outcome if timeouts are not reliable but the network is completely connected and messages arrive within a bounded time? If so, show how. If not, prove consistent outcomes alone will result. (Unreliable means that a site may send an "Are you awake?" message to another site, receive no reply, yet the other site may still be up.)

Question 2

This question has to do with database tuning.

- - For which queries would you expect to get better performance from a non-clustering dense index than from a clustering sparse index? Say why.
  - For which queries would you expect to get better performance from a clustering sparse index than from a clustering dense index? Say why.
- Suppose locking is page-based and your system has a burst of inserts coming from independent transactions into the customer table. The customer table has a sequential key which is customer number and a second non-sequential key which is social security. Which situation would likely have locking problems and how would they compare and why?
  - No clustering index, but a non-clustering index on social security number.
  - Clustering hash index on customer number and a non-clustering index on social security number.
  - Clustering B-tree index on customer number and a non-clustering index on social security number.

Question 3

- Define the consensus problem.
- Say how it relates to the commit problem (if one can be solved, can the other one be?).
- Assuming a reliable ordered broadcast medium with a bounded delay time, but arbitrary timing failures for the processors, give a protocol for consensus.

- In the same setting, is there a non-blocking protocol that guarantees termination of commit (non-blocking means that working processors can decide on how to terminate without waiting for the failed ones to recover)?

Question 4

An airplane control system has three periodic tasks. The first periodic task T1 has a period of 1 second and computation time of .4 seconds. The second periodic task T2 has a period of .7 seconds and a computation time of .2 seconds. The third periodic task T3 has a period of .5 seconds and a computation time of .1 second. Can rate monotonic schedule these tasks? If so, draw out the schedule for the first seven seconds assuming all tasks begin at time 0.

Question 5

Baskerhound at the Post Office.

> "Imagine that a network works like the paper mail service (sometimes called snail mail)," said the chief detective. "Most letters arrive in one day. However, not all do. Some take many days, though they will arrive eventually. The letters may not arrive in order. That is, if A sends a letter to B on day 1 and then another letter to B on day 2, the first letter may arrive last.
>
> "Now, there are 17 people, each of whom has a single message that he or she must send to the other 16 people. At the end, everyone should have the messages from themselves and everyone else. Each letter can contain only one message. All letters must go through a single central post office. Each person has 17 envelopes (one more than may seem strictly necessary up to this point).
>
> "All would be well except that Baskerhound, our adversary, may try to read these messages. He may look at all the letters in the post office at one, perhaps two points of time that he can choose. He won't alter the letters, but he may read them.
>
> "We want to avoid letting Baskerhound know the message held by all 17 people."
>
> "So you have to make sure that copies of all the messages aren't in the mail at the same time?" Ecco asked.
>
> "Exactly," responded the chief detective. "Here is how the solution might go if there were just two people, Bob and Alice and Baskerhound could just look once at the post office. Bob sends to Alice; when she receives, she sends her message to Bob."
>
> "Can we assume that some participant, say number 1, starts the protocol?" asked Ecco.
>
> "Yes, assume that 1 knows when the protocol is to begin," responded the chief detective.
>
> "Finally," Ecco continued, "can a participant send any other message in a letter, e.g. 'I got a message from these participants'?"
>
> "The only other message they can send is the simple phrase 'all clear,' " responded the chief detective.

Answer the following questions ('a' through 'd').

a. Baskerhound can look in the post office only once. We want a solution that will take three or fewer days if all letters arrive in a day, but which doesn't allow Baskerhound to see all the messages no matter how long any given letter takes. Remember that each participant has only 17 envelopes and one message is allowed per envelope. (Actually it can be done with 16 envelopes.)

b. Baskerhound can look twice. Try to find a solution that will take only five days if all letters arrive in a day, but which doesn't allow Baskerhound to see all the messages in his two looks no matter how long any given letter takes. Some participants must use the full 17 envelopes for this problem.

c. Baskerhound can look twice and only needs 10 distinct messages (any 10) to do severe damage. Each participant has 35 envelopes. Try to find a solution that takes eight days, if all letters arrive in a day.

d. Baskerhound can look twice and only needs 10 of the messages (any 10) to do severe damage. However, there are now two post offices (A and B) and participants can route letters through either one. Baskerhound can look at the same post office in his two looks or may use one look for one post office and another look for the other post office. Each participant has 35 envelopes. Try to find a solution that takes four days, if all letters arrive in a day.