Distributed Systems

17th April, 2020
Time: 1:30pm to 3:30pm

This pdf contains three sections.
Section A: Submission Instructions
Section B: Set of questions.
Section C: Topic wise questions which you will refer to in order to answer questions of Section B.
Section D: Appendix
Your answer script should have your roll number and batch on top. Please number the sheets as your write to avoid you losing marks due to the sheets not being ordered and hence not making sense

Section A: Submission Instructions

1. The first line of your pdf should show your roll number and batch (Example: Btech CSE).
2. The answers of all three questions are to be submitted by taking a photo of your handwritten sheets and uploading them as a pdf on moodle.
3. The name of your pdf should be as follows:
<RollNumber>_<TopicNumber of question 2>_<TopicNumber of question3>
For example : "20150231_8_3" if my topic number is 8 for question2 and 3 for question3.
Additionally the below form should be filled after submitting the pdf
https://docs.google.com/forms/d/1fN05xZtnwdzYQTWfLs4C_N7eAMg5aigctH00YBydccg/viewform?edit_requested=true&fbzx=371731649925914754

Section B: Questions

The Quiz contains 3 questions with the following marks and type. The quiz is for max 20 marks.
Question 1. Find which questions you are allowed to attempt. Marks: -5 if wrong and 0 if correct.
Question 2. Answer ANY 1 of the choices you are allowed to attempt. Marks: Max 10
Question 3. Time to create the question! Marks: 12

Question 1: (Marks: Max 0, Min -5)

Consider the below set of topics each marked with a topic number.

Topic Number	Topic Name
1	Clocks
2	Global State
3	Mutual Exclusion
4	Deadlock Detection
5	Termination Detection
6	Wave and Traversal Algorithms
7	Routing Algorithms
8	Minimal Spanning Tree
9	Agreement Protocols

A. Consult the above and write out the topic numbers of the 6 topics of interest that you had mentioned in the google form that was sent out to you. List the topics in increasing order of the topic number assigning them serial numbers 1 to 6. That is, out of the topics submitted by you, the topic with the minimum "Topic number" will be assigned serial 1 and the one whose topic number is maximum with be assigned serial 6.
Note: Serial number and topic number are different. Topic number is the number assigned to a topic in the list given above. Serial number is the number of your selected(submitted) topics when sorted in increasing order of topic number.

B. The choices provided to you for this quiz are as below.
Choice 1: Topic with serial number 5 or 6 . Max marks is 10.
Choice 2: Take the last four digits of your roll number. Suppose it is the integer X. The next choice you have is the question with serial number X mod 4 +1. Max marks is 8.

Thus, if the last four digits of your roll number is 0777 then your next choice available is the topic at serial number 0777 mod 4 +1 = 2

Choice 3: Take the last four digits of your roll number. Suppose it is the integer X. The next choice you have is the question with serial number (X mod 4 +2)mod 4. Max marks is 7.

You need to pick ONLY ONE question to answer. Please neatly mention the serial number and the topic name that you will be attempting in question 2 out of the choices available. Note: the max marks awarded will vary based on the choice you make. In case you have forgotten the choices sent by you please refer to this document (Also given in the appendix)

Question 2 (10 marks)
Topic Wise questions have been provided below in Section C. Neatly mention the topic for which you are answering as selected in Question 1 along with the choice number it falls into. Then continue with answering the question assigned to that topic.

Question 3 (12 marks)
A. For this question, among the 6 topics of your interest you can pick any topic except the one you have already used to answer question 2. You are to provide one true or false question with solution. The true/false question should be 'interesting' and at the same time should be limited to the content covered in class by the instructor (not presentations) which brings out the fundamental understanding of a topic. The question you frame could pertain to a special case or any aspect that you feel a student might overlook which trying to understand the topic or something apart from the two. Such an exercise judges the depth of your understanding of a topic.
We have good google skills as you and have done and will do a thorough job of questions available on the net. I would hence encourage you to be creative and sincere. Overlap found in the questions you frame with those of others will automatically rate you lower as the question could be uninteresting or available on the net. Do your best!

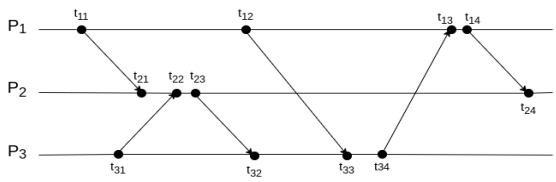
B. How would you grade your own question- interesting, mediocre, uninteresting ?

Section C: Topic wise Questions

Topic 1: Clocks

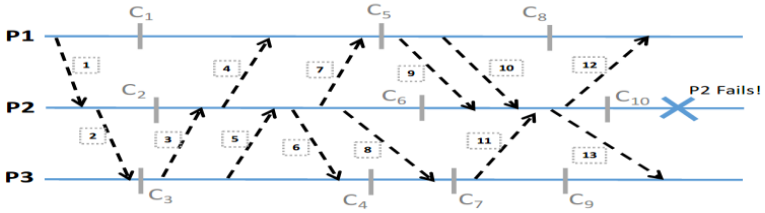
Explain the Singhal–Kshemkalyani's differential technique. What are its advantages and how can the direct implementation be modified to cut down the memory overhead? Are there any prerequisite conditions for this technique to work correctly? Based on the Singhal–Kshemkalyani's differential technique, for the given figure, specify the:

- vector time for each point
- values sent for each message



Topic 2: Global State

You are building a new distributed system where three nodes (P1, P2, P3) communicate by exchanging messages (shown in square boxes). Ideally, you wish to ensure that each node has the same view of the shared state of the system. However, you know that the machines may go down. Nevertheless, the key is to be able to recover from these failures and bring each machine back to a consistent state. To do so, you implement a check pointing solution where each node takes an uncoordinated and independent snapshot of their state of the world (C1 - C10 in the example). Keeping in mind the given scenario, answer the following questions:



A. (8 points) Process P2 fails as shown in the diagram. Can you recover from this failure and get your distributed system back to a consistent state using the different checkpoints you have saved? In particular, calculate a consistent recovery line (listing the snapshots) for rollback. Note:

a. If there is no recovery line state why.

b. If there are multiple, list all of them stating which one is ideal. Explain your choice.

B. (4 points) Explain how you could use logging and replay to augment this checkpointed system to do even better recovery while being more efficient in terms of a checkpoint only solution.

Topic 3: Mutual Exclusion

Consider a parallel setting with the lock server. Multiple threads are in execution which may try to acquire/release the lock. Now, time for some definitions which could be useful:
Mutex: an object that can only be owned by a single thread at any given time
Condition variables: used to wait for a particular condition to become true
Consider the following pseudo-code for lock server:

```
Class LockServer {
private :
        StructureLock {
                ConditionVariable cv ;
                Mutex m;
                Boolean in_use ;
        } ;
        Lock locks[ 6 4 ] ;
        ConditionVariable global_cv ;
        Mutex global_m ;
public:
        void acquire ( LStructure process ) ;
        void release ( LStructure process ) ;
}
```

LStructure is a data structure for storing related information for the lock being considered. Following is the pseudo-code for acquire and release:

```
void LockServer : : acquire ( LStructure lock ) {
        lock the mutex ( global m ) ;
        i f ( locks [ lock . id ] . in use )
                // function atomically release mutex and cause the
                // calling thread to block on the condition variable
                wait_on_condition_variable( global cv , global m ) ;
        locks[lock.id].inuse = True ;
        unlock the mutex ( global m ) ;
}
void LockServer : : release( LStructure lock ) {
        lock the mutex ( global m ) ;
        locks [ lock . id ] . in use = False ;
        // function shall unblock all thread currently
        // blocked on specified condition variable
        broadcast_conditional_variable( global cv , global m ) ;
```

```
        unlock the mutex ( global m ) ;
}
```

Check whether the above code for acquire and release will work or not. If not, give explanation and write pseudo-code for the correct implementation. You explanation should focus on synchronization aspect not memory/time/other irrelevant concepts.

Topic 4: Deadlock Detection

A. Can Phantom deadlocks exist in the Mitchell and Meritts Algorithm. Explain how or why not.
B. Efficiently modify the Chandy-Misra-Haas Algorithm for the AND Model so that when it detects a deadlock, it also knows the lowest priority deadlocked process. Assume each process is given a distinct priority value beforehand.

Topic 5: Termination Detection

Does the algorithm for termination detection by weight throwing, when implemented using a fixed number of bits for storing the floating point numbers, work when the number of rounds and the number of messages can be arbitrarily large? If yes, prove your claim. If no, then find how many bits would be sufficient (need not be minimum) if the number of rounds does not exceed k and the number of processes are n.

Topic 6: Wave and Traversal Algorithms

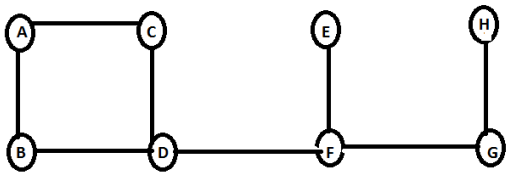
A. In the code below why is the if condition "if $father_p \neq q_0$ AND $\neg used_p[q_0]$" required. Explain this particular if loop by giving an example graph where the code enters this particular if condition.

```
each process, upon receipt of <tok> from q0 :
  begin if father_p = udef then
          begin father_p := q0 ;
              forall r ∈ Neigh_p \ {father_p} do send <vis> to r ;
              forall r ∈ Neigh_p \ {father_p} do receive <vis> from r
          end ;
      if p is the initiator and ∀q ∈ Neigh_p : used_p[q] then decide
      else if ∃q ∈ Neigh_p : (q ≠ father_p ∧ ¬used_p[q])
          then begin if father_p ≠ q0 ∧ ¬used_p[q0]
                       then q := q0
                       else choose q ∈ Neigh_p \ {father_p} with ¬used_p[q] ;
                    used_p[q] := true ; send <tok> to q
               end
          else begin used_p[father_p] := true ;
                    send <tok> to father_p
               end
  end
```

B. Provide a a discussion on the time and message complexity of the Awerbuchs algorithm.

Topic 7: Routing Algorithms

A. Show the running of the Netchange algorithm when the algorithm is applied to a network of following topology. Assume that initially all edge weights are given to be 1.



After a terminal configuration has been reached, a channel between A and F is added. What messages does F send to A when processing the <repair,A> notification? What messages does A send upon receipt of these messages from F ?

B. Can the algorithm be modified to work for graphs where the edge weights are not all 1.

Topic 8: Minimum Spanning Tree

Explain in detail the complexity of the Gallager-Humblet-Spira Algorithm for minimum spanning trees. (5 marks)
How can the algorithm be made modified/made to work on a graph with duplicate(repeated) edge weights. Explain and also show the working of your algorithm on an example graph(the graph must have duplicate edge weights and at least 6 nodes and at least 12 edges).
Can the algorithm also be modified if we assume that we have a situation where both edge weights and node ids, both need not be unique.

Topic 9: Agreement Protocols

We have done the Lamport Shostak Pease Algorithm for the byzantine agreement protocol in class.
A. Take the case of 8 nodes where 2 are faulty. The general can be faulty or non-faulty. Show the running of the algorithm for both cases separately and neatly. Give neat diagrams with step by step explanation as to how the consensus is arrived at by all the non faulty nodes.
B. Will the algorithm work even if general can propagate non binary values.
Note: Do not write out the psuedocode.

Appendix: Link to the choices you have submitted

https://docs.google.com/spreadsheets/d/11M4C7U4dPQrOU_Fyr9Me3_SXIV_FwthLILcg9DlC630/edit#gid=1259579539