# NANYANG TECHNOLOGICAL UNIVERSITY

## SEMESTER 2 EXAMINATION 2020-2021

## CE4013/CZ4013 – DISTRIBUTED SYSTEMS

Apr/May 2021 Time Allowed: 2 hours

## INSTRUCTIONS

1. This paper contains 6 questions and comprises 5 pages.
2. Answer **ALL** questions.
3. This is a closed-book examination.
4. Questions do not carry equal marks.

---

1. (a) What are idempotent and non-idempotent operations?

(3 marks)

(b) A distributed contact tracing system maintains the records of lessons attended by students. The server provides three services. The first service is for the students to register attendance by specifying the location of a lesson (a string), the time of the lesson (two integers, day and hour), and the student name (a string). The second service is for the manager to retrieve all the lessons attended by a particular student by specifying the student name. The third service is for the manager to query the names of all the students attending a particular lesson by specifying the location and the time of the lesson. Design a Java remote interface for the server.

(5 marks)

2. (a) What is the purpose of mounting in a distributed file system? How is mounting carried out in the NFS?

(6 marks)

Note: Question No. 2 continues on Page 2

(b) A distributed file system consists of three computers: a server and two clients A and B. Let *f* be a single-block file stored at the server. Figure Q2 shows the operations performed by clients A and B on file *f*, where "*o*", "*r*", "*u*" and "*c*" represent open, read, update and close operations respectively. A and B both start with an empty client cache, and the caches are large enough to avoid any replacement. The updates made by B do not change the length of file *f*. The transmission delay in the network and the processing times at the computers are negligible.

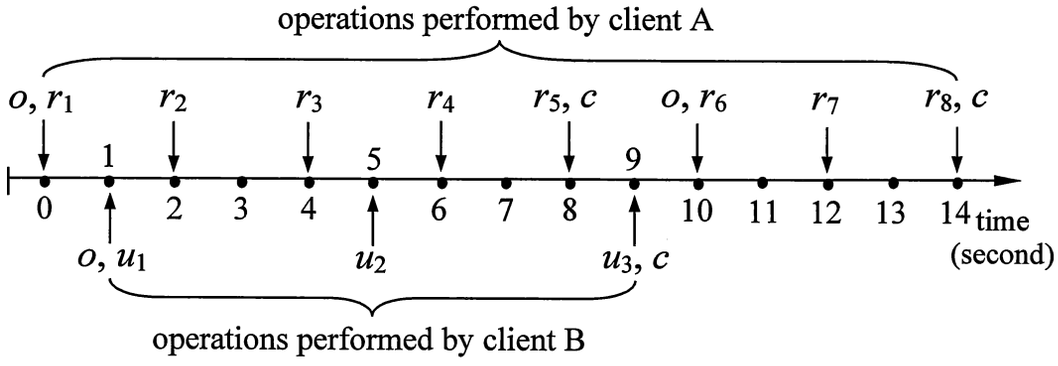


**Figure Q2**

For each of the following cases, determine which read operations of A return the file that incorporates the most recent update by B, and determine the time points when the file is transferred from the server to A.

(i) The system is an NFS. The clients use a freshness interval of 3 seconds to maintain cache consistency, and client B uses a bio-daemon process to send its updates to the server right after making each update.

(7 marks)

(ii) The system is an AFS that implements session update semantics.

(7 marks)

3. Three processes $p_1$, $p_2$ and $p_3$ are running on different computers in an asynchronous distributed system. Each process has a local clock. Figure Q3 shows four messages sent between the processes, where events *a* to *h* represent the sending and receiving events of these messages. Let $t_e$ be the reading of $p_1$'s clock at event *e*, and $t_g$ be the reading of $p_2$'s clock at event *g*. Suppose that $p_3$ knows $t_e$ and $t_g$, as well as the lengths of the time intervals measured by respective clocks as shown in the figure.

Note: Question No. 3 continues on Page 3

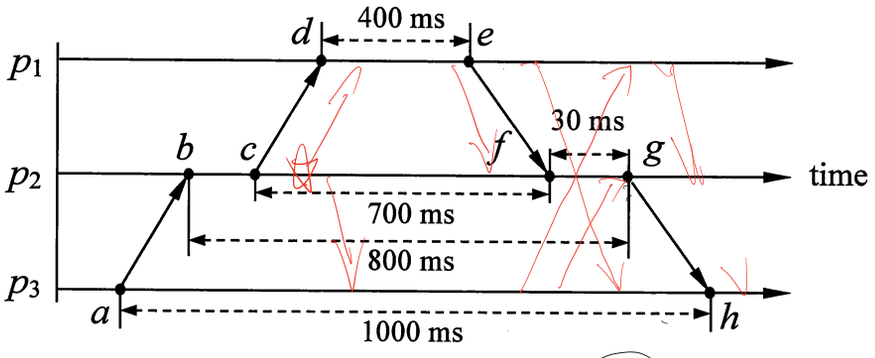


**Figure Q3**

(a) If $p_3$ would like to synchronize its clock with $p_2$'s clock as accurately as possible, what time should $p_3$ set its clock to when it receives the message at event $h$? What is the accuracy of this setting?

(5 marks)

(b) If $p_3$ would like to synchronize its clock with $p_1$'s clock as accurately as possible, what time should $p_3$ set its clock to at event $h$? What is the accuracy of this setting?

(5 marks)

(c) Timestamp all the events in Figure Q3 using the vector clocks.

(6 marks)

(d) Assume that message delivery on each unidirectional point-to-point channel follows FIFO order. $p_2$ initiates the Chandy-and-Lamport algorithm sometime between events $c$ and $f$ to record a snapshot of the system. List all the possible snapshots finally recorded. Indicate clearly the process states and channel states in the snapshots. In your answer, if needed, use $S_1$, $S_2$ and $S_3$ to represent the initial states of $p_1$, $p_2$ and $p_3$ respectively, and use $S_x$ to represent the state of the process where event $x$ occurs immediately after event $x$'s occurrence (for example, $S_a$ is the state of $p_3$ immediately after event $a$ occurs, and $S_c$ is the state of $p_2$ immediately after event $c$ occurs).

(6 marks)

4. The vector clock mechanism is used to timestamp the events that occur in a distributed system of three processes $p_1$, $p_2$, $p_3$. Suppose that in a consistent cut, the last event occurring in each process has the following timestamps:

$p_1$: (300, 100, 100)
$p_2$: (200, 200, 200)
$p_3$: (300, $a$, $b$)

What are the possible ranges of the values $a$ and $b$ respectively? Briefly explain your answer.

(10 marks)

5. (a) Suppose that the Ricart and Agrawala algorithm is used for distributed mutual exclusion among $n$ processes $p_1, p_2, \ldots, p_n$. Below is the partial pseudocode of the algorithm executed by a process $p_i$. Complete the algorithm by filling in the boxes A to F with appropriate pseudocode.

*On initialization*
state := RELEASED;
*To enter the critical section*
state := WANTED **A**
send request to all the other processes;
$T$ := the timestamp of the request by the logical clock;
if (number of replies **B** = number of ru
state := HELD;
*On receipt of a request* $\langle T_j, p_j \rangle$ *from another process* $p_j$ $(j \neq i)$
if (state = RELEASED or state = wanted **C** and $T_j < T_i$)) $i \neq j$
then
Reply immediately to $p_j$ **D**
else
queue request **E** from $p_j$ without reply,
end if
*To exit the critical section*
state := Release **F**
reply to all queued requests;

(10 marks)

(b) In the Ricart and Agrawala algorithm, suppose the statement "$T$ := the timestamp of the request by the logical clock" is changed to "$T$ := a random number". Assume that different processes never generate the same random number. Does the algorithm still guarantee that at most one process may execute in the critical section at any time? Briefly explain your answer.

(10 marks)

6. Consider a replicated shared object service hosting an integer object $x$ whose initial value is 0. Table Q6 shows the operations performed by two clients on $x$. For example, read($x$) $\rightarrow a$ denotes a read operation on $x$ returning a value $a$, and write($x \leftarrow x + 50$) denotes a write operation increasing the value of $x$ by 50. The operations of each client are listed in the order that they are performed by the client.

Note: Question No. 6 continues on Page 5

**Table Q6**

| Program order | Client A | Client B |
|---|---|---|
| ↓ | write($x \leftarrow x + 6$)<br>read($x$) → $a$<br>write($x \leftarrow x + 50$)<br>read($x$) → $b$ | write($x \leftarrow x + 400$)<br>read($x$) → $c$<br>read($x$) → 400<br>read($x$) → $d$ |

(a) If the object service is sequentially consistent, what are the possible combinations of the values $a$ and $b$? (4 marks)

(b) If the object service is monotonic-read consistent, what are the possible combinations of the values $c$ and $d$? (4 marks)

(c) If the object service is both monotonic-read consistent and monotonic-write consistent, what are the possible combinations of the values $c$ and $d$? (4 marks)

(d) If the object service is read-your-writes consistent, what are the possible combinations of the values $a$ and $b$? (4 marks)

(e) If $a$ = 406 and the object service is writes-follow-reads consistent, what are the possible values of $b$? (4 marks)

END OF PAPER

1. 10, 12

2. (2, 12)

3.

**CE4013 DISTRIBUTED SYSTEMS**
**CZ4013 DISTRIBUTED SYSTEMS**

Please read the following instructions carefully:

1. **Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**

2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.

3. Please write your Matriculation Number on the front of the answer book.

4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.