NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 2 EXAMINATION 2021-2022

CE4013/CZ4013 – DISTRIBUTED SYSTEMS

Apr/May 2022 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) In a distributed MRT information system, the server provides a method that allows the clients to query the fare between two specified stations. The server also provides an alert service to the clients through callback. When there is a service interruption for any MRT line, the alert service notifies interested clients of the line name. Assume that the station names and line names are represented by strings, and the fare is represented by a floating-point value.

   (i) Design a Java remote interface for the server to allow the clients to query the fares and to register their interests in the alert service. Design a Java remote interface for the clients to allow the server to perform callback.

   (6 marks)

   (ii) For each method defined in the above interfaces, briefly explain the parameter passing semantics of each parameter.

   (3 marks)

   (b) To resolve a name requested by a client *C*, four name servers are involved in the order of S_1, S_2, S_3, S_4. Suppose that only unicast is used in the name resolution. For each navigation type below, list the sequence of all messages sent among *C*, S_1, S_2, S_3, S_4 in the name resolution process (use the format "$C \rightarrow S_1$" to indicate a message sent by *C* to S_1).

   - Iterative client-controlled navigation
   - Non-recursive server-controlled navigation
   - Recursive server-controlled navigation

   (6 marks)

2. In a peer-to-peer file sharing system based on Chord routing, the identifier circle ranges from 0 to 127 (i.e., $2^7 - 1$). There are 10 nodes in the system: N2, N7, N27, N40, N47, N60, N69, N75, N90 and N96, where the numbers behind 'N' denote the node identifiers on the identifier circle.

   (a) Draw the finger table of node N90.

   (3 marks)

   (b) What is the route of a query issued by node N90 for the location information of file K53 (the number behind 'K' denotes the key identifier of the file on the identifier circle)?

   (4 marks)

   (c) If a query issued by node N7 visits node N40 in its route, what are the possible key identifiers of the target file? (Note: The query may visit other nodes in addition to N40. N40 can be the node responsible for the target file or any other node visited in the query route.)

   (7 marks)

3. Figure Q3 shows three computers A, B and M in an asynchronous distributed system, each having a local clock. M sends requests to the other two computers and then they reply to M. Events a to h represent the sending and receiving events of these messages. Let t_x denote the local clock reading at event x in the computer where x occurs (for example, t_a is the local clock reading of A at event a, and t_c is the local clock reading of M at event c). Assume that M knows the time readings (in units of millisecond) of all events: $t_a = 200$, $t_b = 220$, $t_c = 110$, $t_d = 150$, $t_e = 180$, $t_f = 190$, $t_g = 30$, and $t_h = 60$. The clock drifts at the computers are negligible.

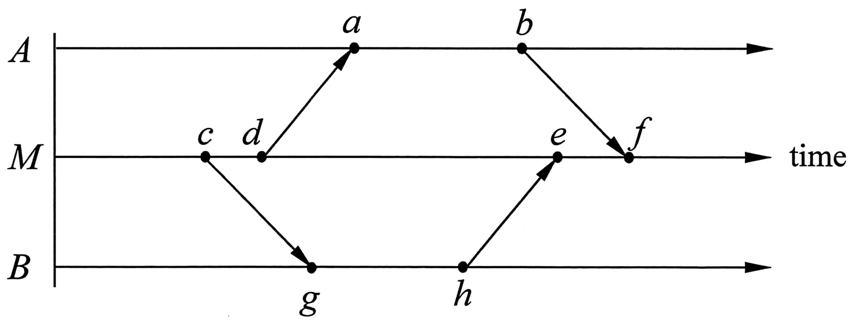

Figure Q3

   (a) M would like to estimate the local clock reading of A. What is M's best estimate of A's clock reading at the time when M receives A's reply (i.e., at event f)? What is the accuracy of this estimate?

   (3 marks)

Note: Question No. 3 continues on Page 3

(b) M would like to estimate the local clock reading of B. What is M's best estimate of B's clock reading at the time when M receives B's reply (i.e., at event e)? What is the accuracy of this estimate?

(3 marks)

(c) Suppose that M acts as the master in the Berkeley algorithm for synchronizing the local clocks of M, A and B. M does not eliminate any clock in its calculation. Following Q3(a) and Q3(b), what time value should M set its clock to when it receives A's reply (i.e., at event f)? What information should M send to A and B respectively for synchronizing their clocks?

(5 marks)

(d) Following Q3(c), after M, A and B perform synchronization using the Berkeley algorithm, what is the largest possible offset between A's clock and B's clock?

(3 marks)

(e) Assume that message delivery on each unidirectional point-to-point channel follows FIFO order. Suppose that M initiates the Chandy-and-Lamport algorithm sometime between events d and e to record a snapshot of the system. What are the possible snapshots finally recorded? In your answer, use S_A, S_B, and S_M to represent the initial states of A, B, and M respectively, and use S_x to represent the state of the computer (where event x occurs) immediately after event x's occurrence (for example, S_a is the state of A immediately after event a occurs, and S_c is the state of M immediately after event c occurs).

(4 marks)

4. Figure Q4 shows the execution of two processes p_1 and p_2 in an asynchronous distributed system, where the dots represent the events occurring in the processes, and the arrows represent the messages sent between the processes. Process p_1 contains a variable x, and process p_2 contains a variable y. The values of x and y at different process states are shown in Figure Q4.

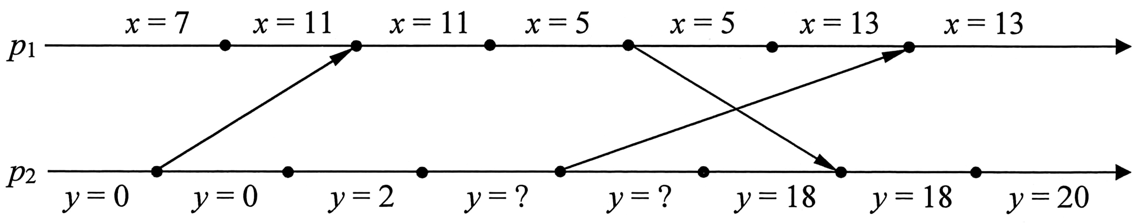

Figure Q4

Note: Question No. 4 continues on Page 4

(a) Draw the lattice of consistent global states. In the lattice, use S_{ij} to denote the global state after i events at process p_1 and j events at process p_2.

(7 marks)

(b) If it can be inferred that the constraint $|x - y| > 6$ must be broken in the execution, what is the possible range of y's values in the two states shown by "$y = ?$" (note that these two states have the same y value)?

(8 marks)

5. (a) We would like to modify the ring-based election algorithm to elect the process with the *smallest* identifier. Assume that each process has a unique identifier. Describe what a process with an identifier 6 needs to do when it receives an election message in each of the following cases:

(i) The identifier in the received message is 5.

(ii) The identifier in the received message is 6.

(iii) The identifier in the received message is 7.

(8 marks)

(b) In the Ricart-and-Agrawala algorithm for mutual exclusion, instead of using logical clock readings to timestamp requests, suppose that every process timestamps its requests with its physical clock readings. Assume that the physical clocks always advance and do not fail. In the case that two requests have the same timestamps, the identifiers of the requesting processes are used to break ties (the process with smaller identifier is given higher priority). With this modification, does the algorithm still guarantee that at most one process may execute in the critical section at any time? Briefly explain your answer.

(12 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($x \leftarrow x + 20$) denotes a write operation increasing the value of x by 20. 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($x \leftarrow x + 100$) read(x) $\rightarrow a$ read(x) $\rightarrow b$	write($x \leftarrow x + 3$) read(x) $\rightarrow c$ write($x \leftarrow x + 20$) read(x) $\rightarrow d$

(a) If $a = 103$ and the object service is sequentially consistent, what are the possible combinations of the values c and d ?

(3 marks)

(b) If the object service is FIFO consistent, what are the possible combinations of the values c and d ?

(3 marks)

(c) If $b = 123$ and the object service provides both monotonic-reads consistency and monotonic-writes consistency, what are the possible values of a ?

(4 marks)

(d) If the object service is read-your-writes consistent, what are the possible combinations of the values c and d ?

(4 marks)

(e) If $c = 103$ and the object service is writes-follow-reads consistent, what are the possible values of d ?

(4 marks)

END OF PAPER

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.