Part II

COMPUTING AND COMMUNICATIONS – Online Assessment [2.5 hrs]

SCC.311 Distributed Systems

---

*Candidates are asked to answer **THREE** questions from **FOUR**; each question is worth a total of 25 marks.*

Question 1

1. Consider a client-server system implemented using Java RMI where multiple clients execute operations (e.g., process data) on a server by sending execute requests. The server is often busy with many requests, and therefore the system designers decided to introduce a **ticketing** mechanism to regulate client requests as explained below.

Before making an execute request, a client must obtain a Ticket object from the server using a getTicket() RMI call. A client must present this ticket along with the execute request. The RMI interface with the two functions is shown below.

```
Public interface BusyServer extends Remote {
        // Get a ticket for the client with the given identifier
        public Ticket getTicket(int 2lientID);
        // Execute operation o with a valid ticket t. Returns false if t is invalid; True otherwise
        public 2lientI execute(Operation o, Ticket t)
}
```

A Ticket object (shown below) returned by the server contains a **validTime** field which contains the *absolute time* (e.g., 12 June 2023 13h:17m:12s) after which the ticket becomes valid; until this time, the ticket is considered invalid by the server. The server only accepts an execute request if it comes with a valid ticket.

Upon obtaining a ticket, the client must wait until the validTime, and then it can immediately send an execute() request to the server, presenting her ticket. The ticket also contains the identifier of the client (ClientID) that is authorised to use the ticket. The server is **stateless** and therefore it does not maintain any per-client state.

```
Public class Ticket implements java.io.Serializable {
   int validTime;    // Absolute time after which this ticket is valid
   int 2lientID;     // Identifier of the client requesting a ticket
}
```

Question 1 continues on next page…

Question 1 continued.

Please answer the following questions about the system described above.

IMPORTANT: In your solutions, you are not expected to provide any code. However, you must mention any new members that you require to be added to the Ticket class and any additions of new RMI calls or modifications to the existing ones that your solution requires. Your solutions for parts a, b, and c below must retain the **stateless** property of the server. You can assume that the clients and the server have synchronised clocks.

1.a. For the above system to work as intended, the server must be able to verify that a ticket supplied by a client in an execute() call was actually generated by the server. Describe how the ticketing mechanism can be extended to achieve this.

[4 marks]

1.b. The clients might be tempted to tamper with the tickets, such as reducing their validTime to execute their operations immediately. Explain how the Ticket objects can be secured against tampering by the server.

[4 marks]

1.c. Propose another extension to the ticketing mechanism so that a ticket is only used by the client that initially obtained the ticket. Explain in detail how your solution works.

[6 marks]

1.d. In this system, each ticket should ideally be used at most once by the authorised client. Explain how the server can achieve this with the minimum amount of state (full marks for a stateless solution).

[2 marks]

1.e. Access to shared Java objects can be serialized by declaring its methods as being synchronized. Is this enough to guarantee serialization when such an object is replicated? Briefly justify your answer. (2-3 sentences)

[3 marks]

Question 1 continues on next page…

Question 1 continued.

1.f. Please select from the following list below the *assumptions* that the Paxos protocol makes about the system model (negative marks for an incorrect selection):

i. Message deliveries can take arbitrarily long times,
ii. Nodes can fail due to crash faults
iii. Nodes can fail due to Byzantine faults
iv. Failure of a node is undistinguishable from delayed messages

[6 marks]

[Total 25 marks]

Question 2

2.a. Consider a collaborative document editing application such as Google Docs where multiple online users can simultaneously edit a common, shared text document that is stored centrally in Google's cloud system. The systems engineers at Google have recently added a new feature to allow users to edit Google Doc documents while offline.

Discuss the challenges of introducing the offline editing mode by referring to an impossibility theorem. (3-5 sentences)

[6 Marks]

2.b. In a system with replicated servers, ten clients broadcast messages to the replicas. The clients stamp their messages with their current system time (at the time of sending the messages). On the other hand, the replicas buffer incoming messages for a fixed amount of time and then process the buffered messages ordered by their timestamps.

i. Assuming that the client messages may take arbitrarily long to deliver, but are eventually received by all the replicas, explain why the approach above would not achieve total ordering.

[2 marks]

ii. If the network guarantees the eventual delivery of each client message at the replicas within a bounded time (e.g., 5 seconds), explain how the above approach can be modified to guarantee total ordering.

[3 marks]

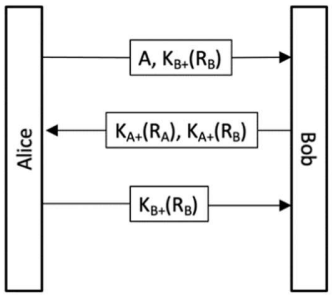
2.c. Is it possible to implement a reliable failure detector using a reliable communication channel? Explain why or why not.

[3 marks]

2.d. Illustrated below is a slightly modified authentication mechanism using public keys. $K_{B+}(X)$ is the encrypted value of X using Bob's public key and similarly, $K_{A+}(X)$ is the encrypted value of X using Alice's public key. A is Alice's identity, R_B is the challenge for Bob, and R_A is the challenge for Alice.

Question 2 continues on next page…

Question 2 continued.



i. Explain how Mallory can exploit a weakness in this mechanism to authenticate herself as Alice using a diagram like the one above.

[5 marks]

ii. Propose a fix to this protocol so that Mallory can no longer authenticate herself as Alice.

[3 marks]

2.e. Name one way of achieving scalability in a distributed system.

[1 mark]

2.f. Discuss why indirect communication is appropriate for mobile environments where network coverage can be poor in some areas. *(2-3 sentences)*

[2 marks]

[Total 25 Marks]

Question 3

Consider the following three protocol pseudocode snippets.

-A-

sender:

```
sendMsg(msg)
for each n in group
send(msg, n)
```

receiver:

```
recv(msg)
deliverToApplication(msg)
```

-B-

sender:

```
sendMsg(msg)
for each n in group except myself
send(msg, n)

for n to maxRetries
sleep(WAIT_TIME)
for each m with no acknowledgement
  send(msg, m)
n ++
```

Question 3 continues on next page…

Question 3 continued.

receiver:

recv(msg)

deliverToApplication(msg)

sendAck(sender, msg)

-C-

sender:

sendMsg(msg)

for each n in group except myself

send(msg, n)

for n to maxRetries

sleep(WAIT_TIME)

for each m with no acknowledgement

  send(msg, m)

n ++

receiver:

recv(msg)

deliverToApplication(msg)

sendAck(sender, msg)

sendMsg(msg)

Question 3 continues on next page…

Question 3 continued.

3.a.

i. Define A:

[1 mark]

ii. Define B:

[1 mark]

iii. Define C:

[1 mark]

3.b. In a network that is generally reliable, how could you improve the efficiency of protocol (B)?

[2 marks]

3.c. Assuming the above protocols are written atop UDP, and one member of a group is sending a single message to all other members of that group, what is the minimum number of messages sent for a group of 6 nodes by protocol B?

[1 mark]

3.d. Explain your answer

[2 marks]

3.e. Using the same assumptions as in the above question, what is the minimum number of messages sent in your modified version of protocol B for a group of 6 nodes?

[2 marks]

3.f. Protocol (C) has an issue with its implementation. What is this issue and how would you correct it?

[3 marks]

3.g. Assuming the above protocols are written atop UDP, and one member of a group is sending a single message to all other members of that group, what is the minimum number of messages sent for a group of 6 nodes by your corrected version of protocol C?

[2 marks]

Question 3 continues on next page…

Question 3 continued.

3.h. Explain your answer

[2 marks]

3.i. Which protocol (A, B, or C) would you use to implement an active replication scheme? [MCQ]

[1 mark]

3.j. Explain your answer with a clear justification for your choice

[2 marks]

3.k Which protocol (A, B, or C) would you use to implement a text-based chat room system which supports many users per room? [MCQ]

[1 mark]

3.l Explain your answer with a clear justification for your choice

[2 marks]

3.m. Which factors of a deployment might we consider in choosing a value for WAIT_TIME in the above protocols? You should consider both the factors that might lead us to choose a higher value and a lower value.

[2 marks]

[Total 25 Marks]

Question 4

4.a. An engineering team has tested the average latency between servers inside their datacentre, which is measured at 1ms. Based on these measurements, the team lead proposes that a failure detector can be built to check on the liveness of each datacentre node by sending a message to that node and waiting 10ms for a response. If no response is received after this time the node will be marked as having failed. Because the response wait period is an order of magnitude greater than the average server latency the team decides it is safe to build software systems that assume their failure detector is perfect.

i. Is this assumption correct? [MCQ]

[1 mark]

ii. Explain your answer with a clear justification

[2 marks]

4.b. Consider the below protocol extracts for a passive replication system, assuming that "processMessage" causes business logic to execute based on the contents of a message:

front end::

messageFromClient(m)

send message m to primary replica

reply to client

replica::

messageReceived(m)

processMessage(m)

reply to sender

send message m to other replicas

Question 4 continues on next page…

Question 4 continued.

i. Is this implementation correct? [MCQ]

[1 mark]

ii. Explain you answer with a clear justification

[2 marks]

4.c. An engineer has been testing the performance of this replication system, and has noticed that the most significant cause of client-observed latency is the sending of messages from the primary replica to backup replicas. The engineer proposes an alternative approach in which the primary replica stores messages in a buffer, and sends them as a batch to backup replicas after every 10 requests. This increases the client-observed speed of the system by almost 8x the original design.

i. Is this alternative implementation correct? [MCQ]

[1 mark]

ii. Explain your answer with a clear justification

[2 marks]

4.d. Consider the below protocol elements for a Byzantine fault-tolerance system:

leader::

onDecision(d)

send d to all members of group G

group-member:

onMessage(d)

add d to message_set for group G

if message_set has one entry for every member of G

decide(majority(message_set))

Question 4 continues on next page…

Question 4 continued.

i. For a system using UDP for message-passing, in a group of 4 nodes (one of which is the leader) the operations management team notices that sometimes group members never decide on an outcome. Why might this be?

[2 marks]

ii. How would you solve the problem?

[2 marks]

iii. Under the same assumptions, for a group of 8 nodes, the operations team notices that sometimes group members decide on different outcomes. Why might this be?

[2 marks]

iv. What modifications need to be made to solve this problem?

[2 marks]

4.e. For a web-based coding help service, the lead architect would like a solution for fault-tolerance to assure system uptime and data retention across failures. The lead architect has specified that low-latency is the most important criteria in choosing a solution.

i. Which replication scheme would you choose to employ:

[2 marks]

ii. Justify your answer

[2 marks]

4.f. For an airline flight booking service, in which customers can visit a website to buy flight tickets, the chief technical officer has requested a report on the ideal replication scheme to ensure data integrity across failures, and has indicated that low-latency is critical to ensure a positive customer experience.

i. Which replication scheme would you choose for this scenario:

[2 marks]

ii. Justify your answer

[2 marks]

[Total 25 Marks]

---End of Paper---