Name:

CSCI/ECEN 5673: Distributed Systems
Spring 2022
Midterm Exam

This is a 1 hour 15 minutes, closed book, closed notes exam. In addition, you will have five minutes to create a PDF file of your answers and upload on Canvas.

Note that the submission link on Canvas will close at 12:25 PM and no late submissions will be accepted.

There are seven questions. You may write your answers on a separate sheet, but do not use more space than what's designated for each question here.

Name:

1. **[20 Points]** Choose one correct answer for each of the following:
   (+2 for correct answer, -1 for incorrect answer, 0 for no answer):

a) An RPC server skeleton (stub):

(i) receives requests from clients and calls the local function on the server.
(ii) is an automatically-generated template for writing server functions.
(iii) is used to discover remote procedures that reside on the server.
(iv) is called when the server-side function cannot be found.

b) Marshaling is the process of:

(i) Converting parameters into a network message for a remote procedure call.
(ii) Sending a message that invokes a remote procedure call.
(iii) The setup process that is needed before any remote procedures can be called.
(iv) The process of a server receiving a message with parameters and calling the appropriate function.

c) An Interface Definition Language (IDL) is used to:

(i) Allow programmers to define server functions in a portable manner so they can run on any system.
(ii) Serialize parameters into a network message.
(iii) List remote functions and their parameters so stubs could be generated.
(iv) Communicate with the network interface to send and receive messages.

d) In Paxos, once an acceptor responds to a Prepare request, it will:

(i) Not accept proposals with higher sequence numbers.
(ii) Not accept proposals with lower sequence numbers.
(iii) Not accept any other incoming proposals.
(iv) Accept any future proposals, regardless of their number.

e) One way in which Raft differs from Paxos is that:

(i) There is no need for majorities.
(ii) In some cases, the resultant value might not be one of the proposed values.
(iii) A single leader to receive all client requests is a requirement.
(iv) The protocol might fail to achieve consensus and will need to be restarted.

Name:

f) When Raft servers hold an election, the winner is generally:

(i) The server with the highest process number.
(ii) The server that picks the highest random number.
(iii) Chosen by a leader, who propagates the choice to a majority of followers.
(iv) The server where a majority of the group members receive its election message first.

g) To survive the simultaneous failure of P acceptors, Paxos requires:

(i) P^2 acceptors.
(ii) P+1 acceptors.
(iii) 2P acceptors.
(iv) 2P+1 acceptors.

h) In an electronic marketplace system, such as the one you are implementing as part of your programming assignments, a safety condition could be

(i) you will never double pay for an item that you purchase.
(ii) you will never have an item in your cart that is not available anymore.
(iii) both (i) and (ii)
(iv) (i) but never (ii)

i) In an electronic marketplace system, such as the one you are implementing as part of your programming assignments, a liveness condition could be

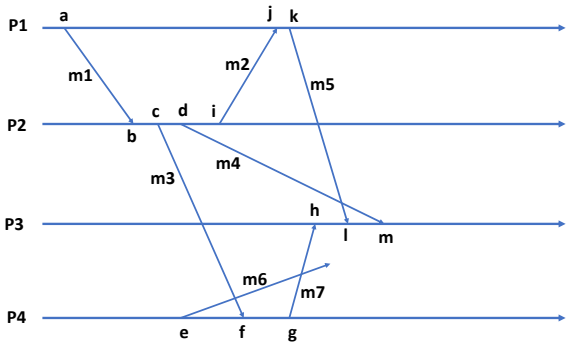
(i) you will eventually be able to purchase an item that is available.
(ii) an item that is not listed for sale by you will never appear in the list of items put up for sale by you.
(iii) both (i) and (ii)
(iv) (i) but never (ii)

j) A fundamental limitation of an asynchronous distributed system is

(i) It is not always possible to reason about the temporal order of events in a distributed system.
(ii) It is not always possible to collect an up-to-date state of the entire system
(iii) both (i) and (ii)
(iv) (i) but not (ii)

Name:

2. Consider the following figure of a message passing system with four processes (**P1**, **P2**, **P3** and **P4**), events (**a**, **b**, … **m**) and messages (**m1**, **m2**, …, **m7**). The system may incur omission communication failures.



Answer the following questions.

a) **[5 Points]** Construct the logical clock (C) values of each event.

b) **[2 Points]** Identify two events *ai* and *aj* to show that C(*ai*) < C(*aj*) does not necessarily imply *ai* → *aj*.

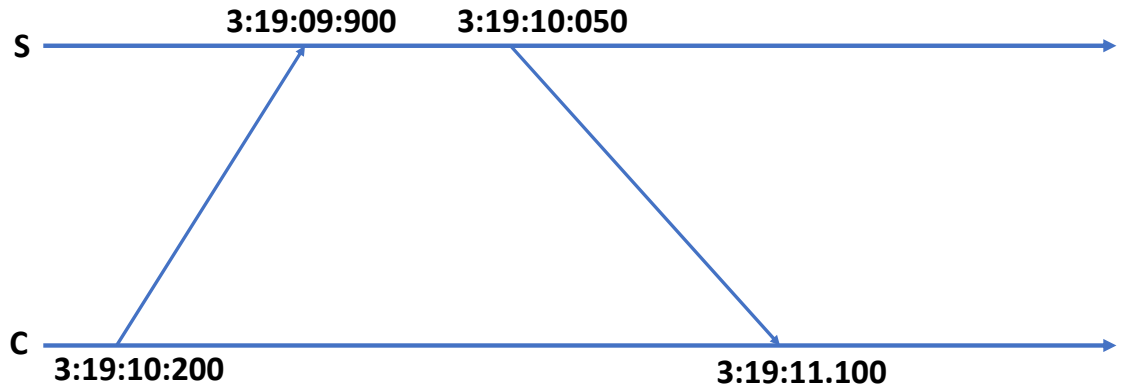
c) **[2 Points]** Is there a violation of FIFO ordering of messages in this figure? If yes, which pair of messages violate FIFO ordering?

d) **[2 Points]** Is there a violation of causal ordering of messages in this figure? If yes, explain where the causal ordering is violated?

e) **[2 Points]** A middleware layer implements total ordering of message delivery that does not violate causal ordering. Delivery of which messages will be delayed?

Name:

3. The local clock of a client (C) is synchronized with a time server (S)'s clock using the NTP. The timestamps are in the format: hour:minute:second:millisecond.



Answer the following questions.

a) **[4 Points]** What is the measure of delay that the NTP client at C will calculate? What are the error bounds in this delay measurement?

b) **[4 Points]** What is the measure of offset that the NTP client at C will calculate? What are the error bounds in this offset measurement?

c) **[2 Points]** For what value of one-way delay would the estimated offset be most accurate?

d) Suppose the client chooses to use Christian's algorithm to synchronize their clock and the server time sent in the response message is 3:19:09:975.

(i) **[3 Points]** What would be the client's estimate of the server's clock value on receiving the response message?

(ii) **[3 Points]** What action the client would take to synchronize their clock after estimating the server's clock value?

Name:

4. **[10 Points]** Consider a group communication system that executes as follows.
   - Current configuration C1 is comprised of members P1, P2, P3, P4 and P5, and messages m1, m2, m3, m4 and m5 are sent in this configuration.
   - Later, due to a network partition and new members joining in, two new configurations, C2 and C3 are installed in parallel.
   - C2 is comprised of members P1, P2, P3 and P6, and messages m6 and m7 are sent in this configuration.
   - C3 is comprised of members P4, P5 and P7, and messages m8 and m9 are sent in C3.

Answer the following multiple-choice questions.
(+2 points for correct answer, -1 for incorrect answer, 0 points for no answer).
In parts (a) to (d), assume that messages are delivered in the order m1, m2, …, m9 by processes that deliver them.

a) Processes P1, P2 and P3 deliver m1, m2, m3, m4 and m5 in C1, and m6 and m7 in C2 along with P6; Processes P4 and P5 deliver m1, m2, m3, m4 and m5 in C1, and m8 and m9 in C3 along with P7.

   (i) This violates extended virtual synchrony but not virtual synchrony.
   (ii) This violates virtual synchrony but not extended virtual synchrony.
   (iii) This violates virtual synchrony and extended virtual synchrony.
   (iv) This does not violate virtual or extended virtual synchrony.

b) Processes P1, P2 and P3 deliver m1, m2 and m3 in C1, and m4, m5, m6 and m7 in C2 along with P6; Processes P4 and P5 deliver m1, m2 and m3 in C1, and m4, m5, m8 and m9 in C3 along with P7.

   (i) This violates extended virtual synchrony but not virtual synchrony.
   (ii) This violates virtual synchrony but not extended virtual synchrony.
   (iii) This violates virtual synchrony and extended virtual synchrony.
   (iv) This does not violate virtual or extended virtual synchrony.

c) Processes P1, P2 and P3 deliver m1, m2, m3, m4 and m5 in C1, and m6 and m7 in C2 along with P6; Processes P4 and P5 deliver m1, m2 and m3 in C1, and m8 and m9 in C3 along with P7.

   (i) This violates extended virtual synchrony but not virtual synchrony.
   (ii) This violates virtual synchrony but not extended virtual synchrony.
   (iii) This violates virtual synchrony and extended virtual synchrony.
   (iv) This does not violate virtual or extended virtual synchrony.

Name:

d) Processes P1, P2 and P3 deliver m1, m2, m3, m4 and m5 in C1, and m6 and m7 in C2 along with P6; and P4 and P5 deliver m1, m2 and m3 in C1, and m4, m5, m8 and m9 in C3 along with P7.

(i) This violates extended virtual synchrony but not virtual synchrony.
(ii) This violates virtual synchrony but not extended virtual synchrony.
(iii) This violates virtual synchrony and extended virtual synchrony.
(iv) This does not violate virtual or extended virtual synchrony.

e) Processes P1, P2 and P3 deliver m1, m2 and m3 in C1, and m4, m5, m6 and m7 in C2 along with P6; Processes P4 and P5 deliver m1, m2 and m3 in C1, and m4, m5, m8 and m9 in C3 along with P7 (m5 is delivered before m4 by all three processes in this case).

(i) This violates extended virtual synchrony but not virtual synchrony.
(ii) This violates virtual synchrony but not extended virtual synchrony.
(iii) This violates virtual synchrony and extended virtual synchrony.
(iv) This does not violate virtual or extended virtual synchrony.

Name:

5. (a) **[3 Points]** Both ABCAST and Totem provide total order delivery of messages in a group in the presence of process and communication failures. Provide one scenario under which you will prefer ABCAST over Totem, and another scenario under which you will prefer Totem over ABCAST.

(b) **[3 Points]** When does a Proposer have to change the value it is proposing in Paxos Consensus protocol?

(c) **[3 Points]** Is it possible in Raft that a leader may not be elected at all? If yes, provide a scenario. If not, explain why not.

(d) **[3 Points]** Why are random numbers used in Raft to set timeout values during leader election?

Name:

6. In the two-phase commit protocol, a network partition causes some (non-zero) participating sites to be partitioned from the coordinator and other participating sites.

a) **[3 Points]** Provide a scenario under which sites in one partition commit while the sites in the other partition abort the transaction?

b) **[3 Points]** Provide a scenario under which sites in one partition commit while the sites in the other partition block?

c) **[3 Points]** Provide a scenario under which sites in both partitions block?

Name:

7. **[20 Points]** A consensus algorithm works as follows. Each process has an initial value, 0 or 1. All n processes ($n > 10$) are organized in a logical ring and the algorithm proceeds in successive rounds as follows:

   - Each process i maintains a vector V_i of size n; initially $V_i[k] = -1$ for all k except i and $V_i[i] = i$'s initial value.

   - In each round, each process i sends V_i to its two neighbors (left and right), and updates V_i based on each vector V it receives as follows:
     if $V_i[k] = -1$ and $V[k]$!= -1, then $V_i[k] = V[k]$ for all k.

   - A process i decides the consensus value to be 0 if more $n/2$ entries in V_i are 0; it decides consensus value to be 1 if at least $\lceil n/2 \rceil$ entries in V_i are 1; otherwise it cannot decide on a consensus value.

   For each of the synchrony and failure models, answer the following two questions: (i) Is consensus among correct processes guaranteed? (ii) When does the algorithm terminate (after how many rounds)? Provide a 1-line explanation for each of your

   (a) Asynchronous system with no process or communication failure.

   (i)

   (ii)

   (b) Asynchronous system, no communication failures, up to 1 process may suffer crash failure.

   (i)

   (ii)

   (c) Synchronous system, no communication failures and up to f ($1 < f < n$) processes may suffer crash failures.

   (i)

   (ii)

Name:

(d) Synchronous system, no process may fail, communication system may suffer message omission failures.

(i)

(ii)

(e) Synchronous system, no communication failures, up to 2 processes may suffer Byzantine failures.

(i)

(ii)