Carnegie Mellon
Computer Science Department.

15-446 Spring 2009
Midterm

Name: _____

Andrew ID: _____

INSTRUCTIONS:

There are 13 pages (numbered at the bottom). Make sure you have all of them.

Please write your name on this cover and at the top of each page in this booklet **except the last.**

If you find a question ambiguous, be sure to write down any assumptions you make.

It is better to partially answer a question than to not attempt it at all.

Be clear and concise. Limit your answers to the space provided.

A	B	C	D	E	F	G
/ 35	/ 12	/ 15	/ 10	/ 10	/ 14	/ 4

TOTAL: / 100

A Multiple Choice - MARK ALL THAT APPLY

1. We discussed two different routing techniques: link state routing (LS), and distance vector routing (DV). Please answer the following questions by circling the protocol(s) for which the claim applies:

   LS, DV - Requires a map of the complete topology

   Solution: -2pts per error LS

   LS, DV - Sends its routing table to its neighbors

   Solution: -2pts per error DV

   LS, DV - Requires flooding

   Solution: -2pts per error LS

   LS, DV - Suffers the count to infinity problem

   Solution: -2pts per error DV

2. The IP stacks of which of the given entities **MUST** be changed for Mobile IP:
   A. Mobile host.
   B. Correspondent host.
   C. Home agent.
   D. All network routers along the path between the mobile host and home agent.

   Solution: +3pts each, -3pts per error
   (a), (c)

A.

P1	W(x)a				
P2		R(x)a	W(x)b		
P3				R(x)b	R(x)a
P4				R(x)a	R(x)b

B.

P1	W(x)a				
P2		R(x)a	W(x)b		
P3				R(x)a	R(x)b
P4				R(x)a	R(x)b

C.

P1	W(x)a			
P2		W(x)b		
P3			R(x)b	R(x)a
P4			R(x)a	R(x)b

D.

P1	W(x)a			
P2		W(x)b		
P3			R(x)a	R(x)b
P4			R(x)a	R(x)b

E.

P1	W(x)a			
P2		W(x)b		
P3			R(x)a	R(x)b
P4			R(x)a	R(x)a

F.

P1	W(x)a			
P2		W(x)b		
P3			R(x)b	R(x)b
P4			R(x)b	R(x)b

3. Which of the above choices could occur on a causally consistent store?

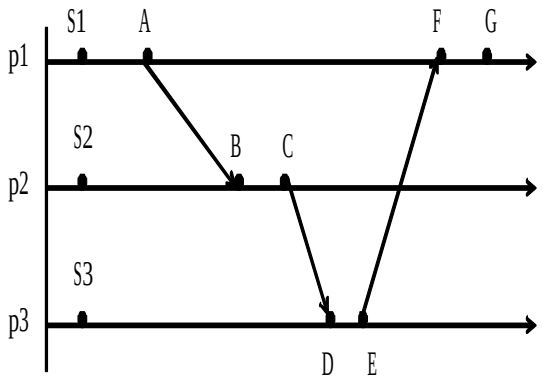
Solution: 1pt each b c d e f

4. Which of the above choices could occur on a sequentially consistent store?

Solution: 1pt each, 1pt extra for all 4 b d e f

5. Which of the above choices could occur on a strictly consistent store?

Solution: 5pt for f, -2pt per extra f



Processes 1, 2, 3 are using a logical clock to keep a consistent time. The horizontal arrow represents physical time and the crossing arrow represents a message being passed. Each point on the horizontal lines is an event.

6. (Lamport Clock) Andrew finds out that the processes above are using Lamport's logical clock, but not all the clock values are not known to him. The initial logical time S1, S2 and S3 were 11, 1 and 0 respectively, and Andrew observed that the clock was 16 at E.

   What is the value of the clock at A, B, C, D, and F?

   Solution: 1pt each, -2 per error

   A=12 B=13 C=14 D=15 F=17

7. (Vector Clock) This time, Andrew finds out that the processes above are using vector clocks, but not all the clock values are not known. The initial logical time S1, S2 and S3 were (0,0,0) , (0,1,0) and (0,0,11) respectively.

   What's the clock values for A, B, C, D, E and F?

   Solution: 1pt each, -1 per error

   A=(1,0,0) B=(1,2,0) C=(1,3,0), D=(1,3,12), E=(1,3,13), F=(2,3,13)

B Short Answer

8. How do the semantics of Birrell's remote procedure call differ from local procedure call? (List at least two ways in which they differ.)

Solution: 3pt each, 1 pt for saying stub instead

Call by value rather than call by reference. Don't pass elements of big structures. Enable call by reference by building a very fast, low interference remote read/write interface like RDMA.

Marshaling by type. With local procedure calls and call by reference, only the formal arguments need to be explicitly typed. But with call by value and complex structures, all of the types of variables in the structures must be known to the marshaling code. This is usually the responsibility of the programmer; using an interface definition language, all the types that must be marshaled are disclosed to the stub generating code.

Performance difference

OS/Language differences

Failure semantics. Local procedure caller and callee failure or survive together. Remote procedure callees and callers fail separately. Servers, the callee, usually try to be stateless, so caller failures don't matter to them, and in some cases provide reply caches to repeat a return result if the caller restarts and retries a call that is not idempotent. Callers often poll callees, usually by retrying the whole operation, unless the operation is idempotent, in which case they may try to poll for "alive and making progress" before invoking a recovery operation (which might be as severe as failing the caller to simulate the "fails together" semantics).

We expect at least two of these to get full marks.

9. Give one benefit and one problem that the programmers obtain by using eventual consistency over sequential consistency.

   (a) Benefit (1-2 sentence answer).

   Solution: 3pts Supports disconnected operation. Sequential systems tend to block operation when correctness cannot be ensured.

   (b) Problem (1-2 sentence answer).

   Solution: 3pts Can result in conflicts and inconsistency. Programmer must deal with this by adding merge techniques or other mechanisms.

C Consistency Implementations

Andrew is given the job of building a distributed file system for mobile hosts. Clients are connected by a low bandwidth, high latency wireless link. Each client has a copy of the entire file system. Andrew needs to make sure that the file system provides a sequentially consistent view of the data to the user. Your job is to use your 446 experience to help him out in designing the system.

10. Andrew decides to use totally ordered multicast to make sure that all nodes apply updates in the same order on all nodes. Unfortunately, this design did not work well as more nodes were added to the system. Give one reason that the system not handling large numbers of nodes well.

> **Solution: 5pts**
>
> When all the mobile nodes are connected to the network, there is still broadcast transmission of the updates and broadcast of the acks to all nodes.
>
> The system also fails to make progress when even one of the mobile nodes is unavailable. This is the main reason.

11. In hopes of improving the system, Andrew switches over to using a quorum-based design. For the following, assume that Andrew chooses the minimum sized read or write quorum sets to ensure correctness. Assume that there are 20 nodes total.
    (a) Nodes read from any single node. How many nodes must they write to ensure correctness?

    > **Solution: 2pts**
    > 20

    (b) Give one advantage and one disadvantage of the above design.

    > **Solution: 1pt**
    > fast reads, slow and unavailable writes

    (c) Nodes write to 11 nodes. How many nodes must they read from ensure correctness?

    > **Solution: 2pt**
    > 10 is actually enough to ensure intersection - but 11 is a fine answer

    (d) Give one advantage of the above design.

    > **Solution: 1pt**
    > highly available

12. With the quorum designs above, users complain that they are not able to update their own files when they are away on travel and disconnected from the network. What alternative design for sequential consistency might you use that allows for disconnected users to perform updates? (hint: think about how Bayou does it – note that Bayou doesn't give sequential consistency though)

> **Solution: 4pt**
>
> Use a primary and make the primary for the files, the owner's laptop
>
> Lots of people picked other parts of Bayou – e.g. use eventual consistency, gossip, vector clocks....

D Following Orders!

Andrew implements a totally ordered multicast system as follows:

1. All nodes uses Lamport timestamps.
2. When node A wants to send a message, it sends it out using TCP to all nodes (including itself). Because we use TCP, you may assume that no messages are lost and that all messages are received in-order.
3. Each message carries the Lamport clock timestamp of the sender.
4. When a node B receives a message, it inserts it into its delivery queue in timestamp order and sends an acknowledgement back to A with its local timestamp (using the same TCP connection).
5. When a node B receives a message, it updates its Lamport clock as appropriate.
6. Whenever a node B wishes to process a message, it takes the first message from its sorted delivery queue.

Andrew believes that this scheme should cause all nodes to process messages in exactly the same order. However, he notices that it does not work in practice.

13. The following table shows the first two events in a sequence. Fill in additional events to show a scenario where node A and node B process M_A and M_B in different orders. Please use the following event types: send, recv+ack+enq, recv_ack, process. Note

Node A	Node B
send M_A, time=4	
	send M_B, time=5

Solution: 6pts

Node A	Node B
send M_A, time=4	
	send M_B, time=5
recv+ack+enq M_A	recv+ack+enq M_B
recv+ack+enq M_B	process M_B
process M_A	recv+ack+enq M_A
	process M_A

14. How would you change Andrew's protocol to ensure that a common order was used on all nodes.

Solution: 4pts

Key missing step is ensuring that everyone has the ack of the message at the head of the queue – this makes sure that it is truly the lowest timestamp..

Therefore, 1) ack must be sent to everyone and 2) process step must block if message is not acked by all.

E Crypto Fun

15. Right before the final, Srini decides that he's had it with Pittsburgh weather and transfers to CMU-Qatar, taking the only copy of the final with him. Dongsu, being the grad student, is stuck with running the class, and not wanting to re-write the whole thing, needs to get the final from Srini in a secure way. Using advanced gnome agents, a group of nefarious students has somehow installed a device that can intercept and modify packets sent from Srini to Dongsu. The exam must not fall into undergraduate hands, or be replaced by a false exam!

    Dongsu and Srini already have both PGP keys and other private keys for each other. These keys have not been compromised.

    For each of the following methods, can the exam be stolen? Can it be replaced with another? Why?

    (a) Srini sends the exam in plaintext.

    Solution: 2pts The exam can be stolen/replaced - no protection in place.

    (b) Srini signs the exam with his private key, sends it.

    Solution: 2pts The exam can be stolen - no encryption. Signing prevents replacement.

    (c) Srini encrypts the exam with Dongsu's public key, sends it.

    Solution: 2pts Can be replaced - everyone has Dongsu's public key. Encryption prevents stealing.

    (d) Srini encrypts the exam with Dongsu's public key, signs it with his private key.

    Solution: 2pts Works...

    (e) Srini signs and encrypts the exam with a shared secret key.

    Solution: 2pts Works...

F DNS Redirection

Harry Bovik is working on a web site that has multiple replicated servers located throughout the Internet. He plans on using DNS to help direct clients to their nearest server replica. He comes up with a hierarchical scheme. Harry has divided his server replicas into three groups (east, west and central) based on their physical location. A typical query occurs as follows:

- When a client makes a query for `www.distributed.hb.com`, the root and `.com` name servers are contacted first. It returns the name server (NS) record for `ns1.hb.com`. The TTL of this record is set to 1 day.
- The `ns1.hb.com` name server is then queried for the address. It examines the source of the name query and returns a NS record for one of {`east-ns, central-ns, west-ns`}`.distributed.com`. The choice of which name server is based on where `ns1` thinks the query came from.
- Finally, one of {`east-ns, central-ns, west-ns`}`.distributed.com`. is contacted and it returns an address (A) record for the most lightly loaded server in its region.

Answer the following 3 questions based on this design.

16. Harry's name server software has only two choices for TTL settings for A and NS records - 1 day and 1 minute. Harry chooses the following TTLs for each record below:

    1. NS record for {`east-ns, central-ns, west-ns`}`.distributed.com` - 1 day TTL.
    2. A record for {`east-ns, central-ns, west-ns`}`.distributed.com` - 1 day TTL.
    3. A record returned for the actual Web server - 1 minute TTL.

    Briefly explain why Harry's choices are reasonable, or why you would have made different choices.

    Solution: 4 points. The name server for a client is based on the region and hence probably does not change very often. Therefore, Harry sets the NS and A records for the name server address to 1 day.

    Harry wants the name server to direct clients to lightly loaded web servers. To do this, Harry must be able to control which web servers each client goes to. Therefore, Harry sets the TTL of the web server A record to 1 minute, so that clients won't cache the record for very long and will ask the name server which web server to use for subsequent requests. If the TTL was 1 day, each client would cache the first A record it got and then continue using the same web server for the entire day even if it becomes overloaded.

17. In general, name resolution systems map names based on the name and context. In this particular case, what are ***TWO*** items of context that the name resolution uses?

Solution: 5pts

1) the IP address of the local name server

2) the load on the servers in the region

18. Harry's Web site is especially popular among CMU students. The CMU network administrator estimates that there is one access from CMU every 3 minutes. Each access results in the application resolving the name `www.distributed.hb.com`. Assume the following:

    - No other DNS queries are made in CMU
    - All CMU clients use the same local name server.
    - This local name server is mapped to the `east-ns` region.
    - Web browsers do not do any caching on their own.

    How many accesses per hour will be made to the following name servers to resolve these CMU queries? Explain your calculation.

    1. The Root Servers
    2. `ns1.hb.com`
    3. `east-ns.distributed.com`

Solution: 5 points.

1. The Root Servers - 1/24 requests/hour
2. `ns1.hb.com` - 1/24 requests/hour
3. `east-ns.distributed.com` - 20 requests/hour

(this page deliberately left blank)

The End – Phew!

G 4 Free Points for Tearing Off Page: Anonymous Feeback

List one thing you liked about the *class* and would like to see more of or see continued (any topic - lectures, homework, projects, bboards, topics covered or not covered, etc., etc.):

List one thing you would like to have changed or have improved about the class: