Department of Electrical Engineering and Computer Science

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.033 Computer Systems Engineering: Spring 2010

Quiz III

There are 13 questions and 11 pages in this quiz booklet. Answer each question according to the instructions given. You have **90 minutes** to answer the questions.

Some questions are harder than others and some questions earn more points than others—you may want to skim all questions before starting.

If you find a question ambiguous, be sure to write down any assumptions you make. **Be neat and legible.** If we can't understand your answer, we can't give you credit!

Write your name in the space below. Write your initials at the bottom of each page.

THIS IS AN OPEN BOOK, OPEN NOTES QUIZ.
You may use a computer to look at PDFs and notes but not for any other purpose.

CIRCLE your recitation section number:

10:00	1. Lampson/Kushman	
11:00	2. Jones/Rieb	3. Rudolph/Kushman
12:00		4. Rudolph/Rieb
1:00	5. Gifford/Post	6. Jones/Spicer
2:00	7. Gifford/Spicer	8. Lampson/Post

Do not write in the boxes below

1-4 (xx/24)	5-9 (xx/30)	10-12 (xx/20)	13-15 (xx/26)	Total (xx/100)

Name:

I Multiple-Choice Questions

1. [6 points]: With respect to the paper "The Recovery Manager of the System R Database Manager" by Gray *et al.* mark each of the following statements true or false.

(Circle True or False for each choice.)

A. True / False Before modifying a "shadowed" file, the entire file is copied, and only the "current" version is modified: the shadowed version is not changed.

B. True / False Before any modified pages can be written to disk, the COMMIT log record for the transaction must be forced to disk.

C. True / False After a checkpoint is written to disk, System R discards all log records that precede the checkpoint record.

D. True / False After saving a shadowed file (making the current version the new shadow version), the old shadow versions of the modified pages can be safely marked as free and reused.

2. [6 points]: In class we learned that DNS is an example of an eventually consistent system. Which of the following statements about DNS are true?

(Circle True or False for each choice.)

A. True / False If DNS is initially configured to resolve name N to IP address A, and is later reconfigured to resolve N to IP address B, clients looking up N after this reconfiguration may continue to receive A as an answer for lookups of N.

B. True / False When there are no network partitions, DNS lookups see changes to DNS records immediately.

C. True / False When consistency has been achieved, a given DNS name resolves to exactly one IP address.

3. [6 points]: Indicate which of the following statements about Ross Anderson's paper "Why Cryptosystems Fail" are true.

(Circle True or False for each choice.)

A. True / False Anderson argues that discussing security failures openly improves security.

B. True / False Most of the security failures described are a result of compromised cryptographic protocols.

C. True / False Anderson suggests that system designers must consider the operation of the equipment as part of the security of the system.

D. True / False The "dual control" concept, where two individuals must collaborate to perform a function, is inconvenient and is sometimes be bypassed by people wanting to save time.

E. True / False Anderson suggests that if we had a set of "perfectly secure" components, a system composed of these components would also be perfectly secure.

4. [6 points]: Indicate which of the following statements about the ObjectStore system described in the paper by Lamb et al (reading 18) are true.

(Circle True or False for each choice.)

A. True / False ObjectStore stores persistent objects in essentially the same format as the ordinary transient C++ objects.

B. True / False To get transactions that are atomic with respect to other transactions that execute concurrently, ObjectStore requires the programmer to lock an object before using it, by explicitly invoking acquire, unlike an ordinary relational database system.

C. True / False If transaction T1 touches objects A and B and no others, and transaction T2 touches objects C and D and no others, then ObjectStore's locking system ensures that T1 and T2 can run concurrently.

D. True / False In ObjectStore an object of type T can be either persistent or transient, and the choice is made when the object is created.

Initials:

5. [6 points]: Indicate which of the following statements about the paper "Hints for Computer System Design" by Lampson (reading 26) are true.

(Circle True or False for each choice.)

A. True / False According to the paper (and the doctrine of 6.033), simplicity of design and interface are always the highest priority.

B. True / False DNS' hierarchical lookups are a good example of following the hint to use brute force.

C. True / False "Keep secrets of the implementation" and "Dont hide power behind an interface" are hints that are often at odds with each other.

D. True / False "Keep basic interfaces stable" is a hint that severely obstructs progress.

6. [6 points]: With respect to the paper "Manageability, availability and performance in Porcupine: a highly scalable, cluster-based mail service" by Saito et. al, which of the following statements is true?

(Circle True or False for each choice.)

A. True / False A given user's mailbox fragments are all stored on the same node.

B. True / False When a failed node recovers after being down for a day the mailbox fragments it stores are brought back up to date from logs on the other nodes.

C. True / False The mailbox fragment list is hard state that keeps track of the nodes that contain a user's mailbox.

D. True / False It is possible that membership services will break a cluster into two disconnected groups of nodes in the presence of certain unusual network failures.

7. [6 points]: Given the context of the paper "Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event" by Kumar et. al which of the following statements are true?
(Circle True or False for each choice.)

A. True / False Witty generated packets with random forged source IP addresses, and thus network telescopes were able to detect replies from victim hosts to the IP addresses fabricated by Witty.

B. True / False The Witty worm exploited a buffer overflow in the ICQ client in Microsoft Windows.

C. True / False It is possible to operate an Internet wide network telescope by telling your computer to listen for the packets addressed to the address range you wish to monitor.

D. True / False The origin of "Patient Zero" was determined by carefully observing which host was first observed at the network telescope.

8. [6 points]: Indicate the truth or falsehood of each of the following with respect to Ken Thompson's paper "Reflections on Trusting Trust."
(Circle True or False for each choice.)

A. True / False Thompson believes that self-reproducing programs shouldn't be trusted.

B. True / False A Trojan horse like the one Thompson describes could not have been hidden in a compiler for a more modern language like Java.

C. True / False The Trojan horse Thompson embedded in the login program could have been found by looking at the machine instructions being executed by the CPU.

D. True / False A programmer can prevent the type of attack Thompson describes by writing all of his or her programs in assembly code.

Buffer overrun, as explained in the paper "Beyond stack smashing: recent advances in exploiting buffer overruns" by Pincus and Baker, can overwrite the procedure return address that is stored on the stack. To avoid this problem, Betty writes a compiler that generates code that maintains two different stacks. One is the usual kind but without the procedure return address, and a second stack, stored in a different part of memory, is used only for the procedure return address.

9. [6 points]: Indicate whether each statement is true or false with respect to Betty's compiler.
(Circle True or False for each choice.)

A. True / False Betty's compiler avoids the classic stack smashing exploit as explained in papers previous to Pincus and Baker, by such hackers as AlephOne and DilDog.

B. True / False Betty's compiler avoids the Arc injection exploit outlined in the paper.

C. True / False Betty's compiler avoids the Function Pointer clobbering subterfuge exploit outlined in the paper.

D. True / False Betty's compiler avoids the Exception-handler hijacking exploit outlined in the paper.

E. True / False Betty's compiler avoids the Heap smashing exploit outlined in the paper.

II Two-Phase Commit

Suppose you are running a transactional 3 node log-based storage system that is using two-phase commit as described in class.

Node 1 is the coordinator, and node 2 and 3 are just workers.

After awhile, node 1 crashes, and the log on its disk looks as follows (here ... indicates some number of UPDATE operations; you can assume in the following three questions that transactions do not UPDATE any of the same records, and that every transaction updates at least one data item):

```
BEGIN T1
BEGIN T2
...
ABORT T1
BEGIN T3
...
COMMIT T3
```

10. [6 points]:

For each of the following transactions, indicate whether the coordinator will end up considering the transaction to have committed or aborted, or whether the information in the log is not sufficient to decide.

(Circle committed, aborted, or can't tell for each transaction.)

A.	T1	Committed	Aborted	Can't tell
B.	T2	Committed	Aborted	Can't tell
C.	T3	Committed	Aborted	Can't tell

Node 1 recovers, and resumes processing transactions. After awhile, node 2 crashes, and the log on its disk looks as follows (assume you know nothing about the coordinator's state other than what is implied by the following):

```
BEGIN T4
...
PREPARE T4
BEGIN T5
BEGIN T6
...
PREPARE T6
...
COMMIT T4
```

Initials:

11. [6 points]:

For each of the following transactions indicate whether node 2 will end up considering the transaction to have committed or aborted, or whether the information in the above log is not sufficient to decide.

(Circle committed, aborted, or can't tell for each transaction.)

A.	T4	Committed	Aborted	Can't tell
B.	T5	Committed	Aborted	Can't tell
C.	T6	Committed	Aborted	Can't tell

Ben Bitdiddle notices that two-phase commit workers have to write two log records for every transaction (a PREPARE record and a COMMIT or ABORT record.) Ben proposes a protocol called *Ben's 2 Phase Commit (B2PC)*. In B2PC, the messages and operation of the protocol are identical to the two-phase commit protocol we learned in class. The only difference is that, when a transaction commits, the workers do not write a COMMIT record to the log (they do, however, still write ABORT records to the log.) When scanning the log during recovery, workers assume that a transaction they prepared actually committed unless an ABORT record for the transaction appears in the log. The coordinator still logs COMMIT records as in the original protocol.

12. [8 points]: Which of the following statements about B2PC protocol are true? Assume that "correct transactional behavior" means the outcome (i.e., COMMIT or ABORT) of the transaction would be the same as in the unmodified 2PC protocol.

(Circle True or False for each choice.)

A. True / False In the absence of crashes or other faults on any of the nodes, B2PC provides correct transactional behavior.

For the following choices, assume that the workers or coordinator may crash, and that there are no other faults in the system.

B. True / False In this case, B2PC provides correct transactional behavior.

C. True / False Assume the coordinator remembers the outcome of all transactions forever. Suppose Ben modifies the protocol described above so that, when recovering from a crash, workers contact the coordinator for the outcome of any prepared transaction that does not have an ABORT record in their log. In this case, B2PC would provide correct transactional behavior.

D. True / False Suppose Ben modifies the protocol described above so that workers do write COMMIT records, but the coordinator omits them. In this case, B2PC would provide correct transactional behavior.

Initials:

III Trusting Ted's Terrific Telegraphic Text Teamware

Theodore is developing a collaborative editor called Ted's Terrific Telegraphic Text Teamware, or TTTTT. Theodore plans to have lots of enthusiastic customers, and he knows that he will have to add new features and issue new releases constantly. He is mulling over the problem of how his customers can verify the authenticity of each new release. A release consists of a single executable file, called ttttt-V.exe (where V is the version number), so the problem boils down to each customer being able to verify that their ttttt-V.exe file has the contents that Theodore intended. Your job is to give Theodore advice about three different authentication schemes he is contemplating.

Theodore's first scheme is to calculate a hash of the content of each release, and to post the hash along with the release on his web site. He uses a cryptographic hash function as described in section 11.2.3 in the course textbook. For each release, Theodore calculates $h_V = H(\text{“TTTTT”} + V + R_V)$, where R_V is the content of the release file ttttt-V.exe and $+$ indicates string concatenation. Theodore does all his development, compiling, and computing of hash values on his laptop, which only he has access to. He posts the release and h_V on his web site at these URLs:

```
http://ted.com/ttttt-V.exe
http://ted.com/V-hash.dat
```

He tells all his customers to fetch both files, and to check for authenticity by comparing the fetched hash value with $H(\text{“TTTTT”} + V + R'_V)$, where R'_V is the contents of the ttttt-V.exe file they fetched. If the two are equal, the customer should accept the new release. If they are not equal, the customer should reject the release.

13. [8 points]: Indicate the truth or falsehood of each of the following statements about Theodore's first scheme.

(Circle True or False for each choice.)

A. True / False An attacker capable of modifying the packets of the customer's transfer of the ttttt-V.exe file could do so in a way way that would cause the customer's ttttt-V.exe file to be different from what Theodore intended, but would cause the customer to nevertheless accept the release.

B. True / False It would be easy for an attacker who can read and modify any files on Theodore's server to cause customers to accept a release that has different contents from what Theodore intended.

C. True / False If a customer sees a “mirror site” of TTTTT that is not affiliated with Theodore, consisting of (for example) http://mirror.com/ttttt-V.exe and http://mirror.com/V-hash.dat, it would be just as safe to fetch those files and compare hashes as it would be to fetch the files from Theodore's web site.

Initials:

Theodore's second scheme is to generate authentication tags for each software version using a shared-secret message authentication code (MAC), as described in the textbook in sections 11.3.3, 11.3.4, and 11.3.5. Theodore generates a separate key K_i for each of his customers, and contacts each customer on the telephone to give them their K_i. For each new release V, Theodore calculates an authentication tag $T_{V,i}$ for each of his customers by calling SIGN("TTTTT" $+ V + R_V$,K_i), where R_V is the content of ttttt-V.exe. Theodore does all his development, compiling, and computing of MAC values on his laptop, and he stores the K_i keys only on his laptop; only Theodore has access to this laptop. Theodore posts the release file and all the tag files on his web site, at these URLs:

```
http://ted.com/ttttt-V.exe
http://ted.com/V-1.tag
http://ted.com/V-2.tag
...
```

He tells his customers to each fetch the new release and their tag from the web site and to call VERIFY("TTTTT"$+$ $V + R'_V$,T,K_i), where R'_V is the content of the ttttt-V.exe file they fetched and T is the content of the V-i.tag file they fetched. If the call returns ACCEPT, the customer should accept the new release; otherwise the customer should reject the new release.

14. [9 points]: Indicate the truth or falsehood of each of the following statements about Theodore's second scheme.

(Circle True or False for each choice.)

A. True / False An attacker capable of modifying the packets of the customer's transfer of the ttttt-V.exe file could do so in a way way that would cause the customer's ttttt-V.exe file to be different from what Theodore intended, but would cause the customer to nevertheless accept the release.

B. True / False It would be easy for an attacker who can read and modify any files on Theodore's server to cause customers to accept a release that has different contents from what Theodore intended.

C. True / False If a customer sees see a "mirror site" of TTTTT that is not affiliated with Theodore, consisting of (for example) http://mirror.com/ttttt-V.exe and http://mirror.com/V-i.tag files, it would just as safe to fetch those files and verify with VERIFY as it would be to fetch the files from Theodore's web site.

Initials:

Theodore's third scheme is to configure his web server to use SSL with a certificate from a well-known certificate authority. Theodore's server holds a particular private key, and the certificate says that whoever knows that private key is the rightful owner of the DNS domain ted.com. Theodore tells his customers to fetch new releases from

```
https://ted.com/ttttt-V.exe
```

Theodore tells his customers that they do not have to take any special steps to check the authenticity of the software, since, if you tell a web browser to connect to https://*servername*, it will use SSL to check that the server can prove ownership of a certificate for *servername* from a well-known authority. Theodore tells his customers to check that the right URL appears in their browsers' URL box. SSL (also known as TLS) was described in lecture, and you can also read about it in section 11.10 of the textbook.

15. [9 points]: Indicate the truth or falsehood of each of the following statements about Theodore's third scheme.

(Circle True or False for each choice.)

A. True / False An attacker capable of modifying the packets of the customer's transfer of the ttttt-V.exe file could do so in a way way that would cause the customer's ttttt-V.exe file to be different from what Theodore intended, but would cause the customer to nevertheless accept the release.

B. True / False It would be easy for an attacker who can read and modify any files on Theodore's server to cause customers to accept a release that has different contents from what Theodore intended.

C. True / False If a customer sees a "mirror site" of TTTTT that is not affiliated with Theodore, consisting of (for example) https://mirror.com/ttttt-V.exe, it would be just as safe to fetch those files via SSL as it would be to fetch the files from Theodore's web site.

End of Quiz III

Please double check that you wrote your name on the front of the quiz,
and circled your recitation section number.

Initials: