CSci 555 Midterm Exam

Fall 2007

Instructions:

Show all work. No electronic devices are allowed. This exam is open book, open notes. You have 120 minutes to complete the exam.

Please prepare your answers on separate sheets of paper. You may write your answers on the sheet of paper with the question (front and back). If you need more space, please attach a separate sheet of paper to the page with the particular question. Do NOT extend your answer on the back of the sheet for a different question, and do NOT use the same extra sheet of paper to answer more than one question.

In particular, each numbered questions must appear on separate pieces of paper so that the exam can be split for grading.

Be sure to include your name and USC ID number on each page.

There are 100 points in all and 4 questions.

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. (30 points) Communication Models

In class we discussed several communication models including message passing, distributed
shared memory, and remote procedure call. This question will test your understanding of the issues
related to these communication models.

a) Explain the difference between message passing and a remote procedure call (in its most basic
form) in terms of the following characteristics: (in answering these questions you might indicate,
where appropriate, what extra work is needed by the programmer using such a mechanisms to
achieve the characteristic). (20 points)

Synchronization:

Parallelism:

Reliability:

Performance:
b) You are developing a distributed database that will operate across a high-latency satellite link. You are considering distributed shared memory and message passing approaches for communicating data between parts of the system. Which approach would you choose and why? Why would you rule out the other approach? [10 points]
2. (20 points) Replication and caching

The use of replication and caching in distributed systems improves performance and reliability, but it requires attention to consistency of data. Use the term strict consistency to require that a read will always return the data that was most recently written. For each of the systems that follow, indicate whether it provides strict consistency (strict) and if so provide a single sentence explaining why, or indicate that it does not provide strict consistency (non-strict) and if not provide a single sentence explaining when/how such inconsistent information is returned.

The Domain Name System:

Quorum Consensus (Weighted Voting)

Grapevine

Transactions (this one may be a little bit tricky)
3. (15 points) Naming

For each example below, explain the kind(s) of naming that is employed in terms of “host-based naming”, “global naming”, “attribute based naming”, and “user, host, or object centered naming”. Note that some examples may employ more than one kind of naming, and if so you should identify the possible overlapping components that use each. Other examples may fall into one category as envisioned, but may in reality fall into another category because of various tricks that are done within the implementation. Explain why you put each name into a particular category.

a) The hostname “pollux.usc.edu”

b) The URL http://gost isi.edu/555
c) The email address “csci555@usc.edu”

d) The filename “/bin/cat” on a Unix system

e) The incomplete X.500 name /C=US/OU=Department of Computer Science/
4. **(35 points) Design problem**

You have been hired as a consultant the new company Actrans that is developing a distributed system to manage transactions for banks and other financial providers. Customers will sign up with different providers, yet they will need to enter into transactions with accounts managed by different providers. You will be responsible for advising the providers on their needs for availability, performance, consistency, and security in the computer systems they deploy and in evaluating how proposed technologies meet those needs.

a. Explain the consistency requirements for such a system (i.e. what are the transaction properties needed by such a system). (10 points)

b. Where would you store the data of customers, including account balances, in order to improve performance and security? How does this placement of data affect the availability of access to balances, and how does it affect the consistency mechanisms that are employed during transactions (you can give multiple example for different scenarios). (10 points).
c. How will you identify the server that manages a customer’s account (i.e. where the balance is stored)? (5 points).

d. Discuss the security requirements for such systems. In particular, how will a transaction be authorized by a user? How will the user be authenticated when they initiate a request to transfer funds from one account to another? How will data be protected when it is resident on disk, and how will communication needed to complete transactions be protected as it flows between participating systems? (10 points)