Name _____

Cpt. S. 464/564 Distributed Systems Concepts and Programming

Final Examination
May 5, 2003
3:10-5:10 PM

80 possible points (464) ____ check here if 464 student
124 possible points (564) ____ check here if 564 student

25% of final grade
8 Pages

For each problem, give your answer in the empty space following the entire problem. If you need more space, use the back of the sheet and clearly mention this on the front. Questions that ask you to *define* a term should be answered succinctly. Questions that ask *why,* or ask you to *explain, discuss, describe,* or *give an example* invite longer answers. If you are not sure about some part of the question, state any assumptions you are making about the question, then answer it based on those assumptions. If the assumptions are reasonable, there will be no points off.

1) [464-10 Points, 564-14 points] Naming, Directory, and Discovery Services

a) [6 points] Define "Naming Service", "Directory Service" and "Discovery Service". Your definitions should draw clear distinctions between the different kinds of service.

b) [4 points] Give two or more reasons why naming services are often organized hierarchically.

c) [4 points; <u>564 only</u>] Describe how the GNS accommodates adding a new root level that encompasses several existing roots, without breaking any existing name references.

2) [464-12 points, 564-18 points] Distributed File Systems

The designers of the Andrew File System observed that files are usually small, that reads were more common than writes and were usually sequential, that most files were updated by a single user, and that when a file was accessed it was likely to be accessed again in the near future.

a) [6 points] Describe the implications of these observations for the design of AFS

b) [6 points 564 only] Pick one of these observations and give an example of a usage pattern that violates it, leading to poor performance of an AFS-based system.

c) [6 points 464&564] One of the claimed design advantages of NFS is that the servers can be stateless. Why is it an advantage? Under what circumstances might it be a disadvantage?

3) [12 points] Network Security Mechanisms

a) [2 points] What is physical circuit encryption?

b) [2 points] What is virtual circuit encryption?

c) [4 points] What are the deployment issues associated with physical circuit encryption?

d) [4 points] What are the deployment issues associated with virtual circuit encryption?

4) [14 points] Real-time (RT) Systems

a) [2 points] Define Hard RT: be sure to include a discussion of assumptions about the environment's behavior.

b) [2 points] Define Soft RT

c) [4 points] What is the relative difficulty of implementing RT systems when work arrives sporadically, aperiodically or periodically? Explain

d) [2 points] Define Real-time communication

e) [4 points] Why is flow control necessary in a network intended to support real-time communication?

5) [12 points] Types of Real-time (RT) Scheduling

Briefly compare and contrast each of the following kinds of RT scheduling. For each pair, indicate which category *rate monotonic scheduling* falls into.

a) [4 points] Static vs. Dynamic Scheduling

b) [4 points] Fixed-Priority vs. Dynamic-Priority Scheduling

c) [4 points] Pre-emptive vs. Non-Preemptive Scheduling

6) [464-10 points 564-20 points] Security

a) [2 points] Define "failure"

b) [2 points] Define "error"

c) [2 points] Define "fault"

d) [2 points] Define "vulnerability"

e) [2 points] Define "threat"

f) [10 points 564 Only] Describe the interplay of all these concepts in assessing risks to distributed systems.

7) [464-12 points 564-22 points] Overlay networks

a) [2 points] The Pastry overlay network is claimed to route messages from source to destination in log N overlay hops, on average. What is N in this expression?

b) [10 points] The Pastry node ID corresponding to a particular internet (IP) address is random. How, then, can a message be sent so quickly to its destination?

c) [10 points 564 Only] Describe how an overlay network such as Pastry can be used to implement a distributed hash table, with replication for availability.

8) [12 points, 564 Only] Order

GNS guarantees its clients that the answers it supplies will be based on all of the updates with timestamp prior to the "last sweep time" and none, some, or all of the updates since then.

a) [6 points 564 Only] Describe the mechanism used to implement this guarantee. (You don't need to describe how joining and leaving servers are handled, but you must address what happens when a server is temporarily down.)

b) [6 points 564 Only] What would be the consequences for the system design if GNS instead tried to provide a guarantee that clients would see updates in causal order: that is, if a client read data from GNS and later submitted an update then no other client could see the update unless they could also see the data that was read prior to the update, regardless of which server the other client happened to use.