Study Sheet for Midterm #2

1. What are the advantages of using a simulation to evaluate the performance of a system?
2. What are the advantages of using a measurement study to evaluate the performance of a system?
3. What are the advantages of a passive measurement study? An active measurement study?
4. What are some ways of performing passive measurements to collect information about system performance?
5. What are some ways of performing active measurements to collect information about system performance?
6. For a given system, be able to explain the topology, workload, and metrics you would use to evaluate the system.
7. Why is it important to model the workload of a system?
8. What is a trace-driven workload, and what are its advantages and disadvantages?
9. What is stress-testing, and when would you use it?
10. What are the advantages of a synthetic workload?
11. Why are a statistical mean or median often not enough to characterize a distribution with high variability?
12. What is a PDF? What is a CDF?
13. What does an exponential or pareto distribution model, that can’t be captured in a uniform distribution?
14. What is Zipf’s Law?
15. Explain how the original Napster service worked.
16. Explain how Gnutella version 0.4 supports peer discovery and content location.
17. Explain how Gnutella version 0.6 uses UltraPeers to provide more scalable peer discovery and content location.
18. Why is peer-to-peer content delivery able to scale to handle higher loads as compared to a single web server?
19. Explain how BitTorrent performs bootstrapping, peer discovery, and content location.
20. Explain how BitTorrent’s incentives work. How does it help new connections to avoid starvation?
21. What are the advantages of designing a web database application as opposed to a traditional stand-alone application?
22. In a relational database, what is a primary key?
23. Give an example of a one-to-many and a many-to-many relationship in a relational database.
24. Explain how you can resolve a many-to-many relationship into two separate one-to-many relationships using an extra table.
25. In entity-relationship modeling, what is a weak entity? What is a full participation relationship?
26. In MySQL, what does the NOT NULL keyword do?

27. What is a JOIN query? How is an INNER JOIN different from an OUTER JOIN?

28. Given a database (such as the one used in the book), be prepared to formulate an SQL query that will answer a given question. For example, show a list of customers that have purchased a particular juice.