

May 2012

98-Comp-B10, Distributed Systems

3 hours duration

NOTES:

1. If doubt exists as to the interpretation of any question, the candidate is urged to submit with the answer paper a clear statement of any assumptions made.
2. This is a CLOSED BOOK examination.
A Casio or Sharp approved calculator is permitted.
3. Answer any **six of the seven** questions.
Only the first five questions as they appear in the answer book will be marked.
4. All questions carry equal weight.
5. Most questions require an answer in essay format. Clarity and organization of the answer are important.

Question # 1. Characteristics of distributed systems

- a. Name two advantages and two disadvantages of distributed systems over centralised systems.
- b. Explain what is meant by administrative scalability and why it is often such a difficult problem to solve.

Question # 2. Fundamental concepts and mechanisms

- a. Name and define two performance characteristics of a communication channel.
- b. Can two or more computers in the Internet have the same IP addresses? Justify your answer
- c. Comment on the use of Connectionless (UDP) and connection-oriented (TCP) communication for of each of the following application:
 1. Virtual terminal access (for example, Telnet);
 2. File transfer (for example, FTP);
 3. User location (for example, rwho, finger);
 4. Information browsing (for example, HTTP);
 5. Remote procedure call.

Question # 3. Client-server systems & inter-process communications

- a. Is it important to consider the number of messages exchanged in a protocol as a performance metric? Or it is enough to consider only the total amount of data sent? Please justify your answer. Design a variant of the RRA protocol in which the acknowledgement is piggy-backed on, that is, transmitted in the same message as, the next request where appropriate, and otherwise sent as a separate message.
- b. Develop a scheme that utilizes message retransmissions with IP multicast to overcome the issue of dropped messages.

Question # 4. Operating systems for distributed architectures

- a. A file server uses caching. The average hit rate achieved is 85%. If the requested block is in the cache, file operations in the server cost 4 ms of CPU time; otherwise, it takes an additional 15 ms of disk I/O time otherwise. Estimate the server's throughput capacity (average requests/sec) if it is: (Explain any assumptions you make)
 1. single-threaded;
 2. two-threaded, running on a single processor;
 3. two-threaded, running on a two-processor computer
- b. Compare the thread-per-request architecture with the worker pool multi-threading architecture.
- c. What the kernel must provide for a user-level implementation of threads (such as Java on UNIX)? In user-level threads implementations, do page faults present a problem?

Question # 5. Security

- a. What is the difference between message confidentiality and message integrity? Can you have confidentiality without integrity? Can you have integrity without confidentiality? Justify your answer.
- b. Pretty Good Privacy (PGP) is broadly used to secure email communication. Explain the steps needed prior to exchanging email messages with privacy and authenticity guarantees for a pair of users using PGP.

Question # 6. Distributed file systems

- a. Name and discuss three key design issues for distributed file systems.
- b. Compare AFS and NFS from stability point view? Are there any limits on AFS scalability, assuming that servers can be added as required? Are there any recent technological developments that would help to offer greater scalability?

Question # 7. Principles of fault tolerance continual

- a. What is the gossip architecture? Why does a replica manager need to keep both a 'replica' timestamp and a 'value' timestamp?
- b. Write pseudocode for a scheme for integrating two replicas of a file system directory that underwent separate updates during disconnected operation.