

# Systems Comprehensive Exam, Spring 2007

January 8, 2007

## 1 Instructions

This is a closed-book, closed-notes exam with a total of 100 points. You may not use any external source for answering these questions, including but not limited to the Internet, books, notes, or other people. Please direct any questions about this exam to Professor Bridges. Professor Bridges may be reached either in person in his office in 301B Farris, by phone at 277-3032 or 363-8798, or by email at [bridges@cs.unm.edu](mailto:bridges@cs.unm.edu). Turn your exam in to Professor Bridges or the front office by 5:00 PM MST Monday, January 8, 2007. Exams *will not* be accepted after this time except by prior arrangement with Professor Bridges.

Type or write your answers to the stated number of questions in each of the following three sections. Make any *reasonable* assumptions necessary to answer the question, but be sure to state any assumptions that you make.

## 2 Short Answer - Answer 3 of 4 (30 points)

*Briefly* answer 3 of the following 4 questions. Your answer should be no longer than *one* paragraph.

1. What is an access control list? Describe the information that goes into an access control list and describe how this information is used.
2. Explain the difference between internal and external fragmentation in a memory system. Which type of fragmentation is associated with variable sized allocation?
3. In the context of job scheduling, describe the difference between throughput and response time.
4. Give a brief description of polling and interrupt driven I/O. Describe a circumstance in which you might prefer polling over interrupt driven I/O.

## 3 Medium Answer - Answer 2 of 3 (40 points)

Provide detailed answers to two of the following three questions. Be sure to state any assumptions you make and to fully justify your answers. Limit your answers to approximately one to two pages in length.

1. Briefly describe each of the ACID properties for a transaction.
2. In a demand page virtual memory system, describe the sequence of steps that the hardware or operating system must perform when a process tries to access a logical address that is not currently mapped for the process.
3. Suppose that a process is processing a file in sequential order. Explain how double buffering will improve the overall performance of the process. Start by describing double buffering. Are there circumstances under which adding a third buffer would be beneficial?

## 4 Design - Answer 1 of 2 (30 points)

Provide a *full* and *detailed* answer to one of the following two questions. Be sure to state any assumptions you make and to fully justify your answer.

### 4.1 Consistency in Data Processing Systems

Consider a general data processing/data mining system where data updates are continually being streamed into an archive, and a backend system is processing the archive for useful information. Specifically, consider the following two examples:

1. A web indexing system used by companies such as Google, where the data updates are web pages crawled by concurrent web spiders, and the backend processes the resulting archived graph of web pages to generate indexes that are used to satisfy search requests.
2. A business database/intelligence system for a web retailer such as amazon, where the data updates are sales orders generated by web servers which include the contents of the order, the customer ID of the person placing the order. The resulting database is used both for:
  - (a) Tracking inventory and order status by web applications
  - (b) Mining suggestions to customers about what to purchase

In each of the example systems above, analyze and describe the demands of and tradeoffs between consistency and performance in this system. Be sure to consider:

- The potential need for consistent snapshots of the data being acquired by the front-end systems;
- The potential need for consistency versus data acquisition performance between multiple writers in the front-end data acquisition systems;
- The potential need for consistency versus data analysis performance between front-end data acquisition writers and back-end data analysis readers.

### 4.2 Diagnosis in Autonomic Systems

Diagnosis of component failure is a critical part of autonomic systems. Describe how the need to run diagnostics may affect the design of systems software and the overall system. Be sure to consider diagnostics that require exclusive use of a component and the degree to which the system needs to be highly available.