Chalmers Un. of Technology and Gothenburg Un. Comp. Science and Engineering Department

Operating Systems
DIT 400, EDA092

Exam 2012-12-17

Date, Time, Place: Monday 17/12 2012, 14:00-18:00, H building

Course Responsible: Arne Dahlberg Tel: 772 1705

Auxiliary material: You may have with you
- An English-Swedish, Swedish-English dictionary.
- No other books, notes, calculators, PDA's etc.

Grade-scale ("Betygsgränser"):
CTH:3:a 30-38 p, 4:a 39-47 p, 5:a 48-60 p
GU: Godkänd 30-47p, Väl godkänd 48-60 p

Instructions

- Do not forget to write your personal number, if you are a GU or CTH student and at which program ("linje").
- Start answering each assignment on a new page; number the pages and use only one side of each sheet of paper.
- Write in a **clear manner** and **motivate** (explain, justify) your answers. If it is not clear what is written, your answer will be considered wrong. If it is not explained/justified, even a correct answer will get **significantly** lower (possibly zero) marking.
- If you make any assumptions in answering any item, do not forget to clearly state what you assume.
- The exam is organized in groups of questions. The credit for each group of questions is mentioned in the beginning of the respective group. Unless otherwise stated, all questions in a group have equal weight.
- Please answer in English, if possible. If you have large difficulty with that and you think that your grade can be affected, feel free to write in Swedish.

Good luck !!!!

1. (10 p)

   (a) Many UNIX commands are able to work as a "filter". What are the requirements on a program to be able to work as a "filter"? (2p)

   (b) Write a simple C-program (shell) that demonstrates how to create a pipe between two processes. The program shall fork two new processes and set up a pipe between these processes. Let one process exec the *ls* command and the other exec the *wc* command. (8p)

2. (10 p)

   (a) Explain how a system call is done at assembly level (2p)

   (b) If a user program reads a file on a local file system or an NFS file system, the same system call will be used in both cases. Explain the mechanism that makes the use of different file systems transparent in file system accesses. (2p)

   (c) What can you expect to find in a UNIX file system directory for a normal file? (1p)

   (d) Why do UNIX hard links not generate cycles in the file system graph? (1p)

   (e) In almost all operating systems the first block in a disk partition is reserved for a special purpose. Which purpose? And is it necessary to reserve the block for this purpose? (2p)

   (f) The BSD filesystem used rotationally optimal placement of data blocks. Newer filesystems like the Linux ext3 filesystem use sequential placement. Which change in hardware have lead to sequential access being preferred today? Explain why. (2p)

3. (10 p)

   (a) The following data is given for a system: Demand paging, with the page table in internal registers. The memory access time is 100 ns. A page fault and the following read operation requires 8 ms plus another 12 ms if a modified page is replaced. Assume that 70% of the replaced pages are modified. What is the maximum page fault frequency that can be accepted if an effective access time of maximum 200 ns is wanted? (4 p)

   (b) Describe the clock algorithm for page replacement. Assume that you are monitoring the rate at which the pointer in the clock algorithm (the one that indicates the candidate page for replacement) moves. What can you say about the system if you notice that it is moving fast? What if it is moving slow? (4 p)

   (c) Consider a logical address space of eight pages of 1024 bytes each, mapped onto a physical memory of 32 frames. i. How many bits are there in the logical address? (1p) ii. How many bits are there in the physical address? (1p)

4. (10 p)

   (a) Explain how code can be shared among different processes in a system that uses virtual memory based on paging. (2p)

   (b) Explain what a page table entry is and the typical information fields it contains. (2p)

   (c) What are the two most important reasons to use RAID systems? (2p)

   (d) What is the most important difference between RAID 5 and RAID 6? (1p)

   (e) Mention two types of errors that RAID do not protect against. (1p)

   (f) Consider the disk scheduling algorithms SCAN and FCFS.
   i. Which of the methods is likely to give best performance for a heavily loaded system? (1p)
   ii. Which of the methods is likely to give best performance for a very lightly loaded system (only one request in the queue)? (1p)

5. (10 p)

   (a) Explain why NFS version 3 cannot use a server-initiated method for cache validation. (1p)

   (b) What does it mean that the naming method in a distributed file system is *location transparent*? (1p)

   (c) What is the main advantage with using *Session Semantics* compared to UNIX semantics? (1p)

   (d) In NFS, an encoding called XDR (External Data Representation) is used for data sent between client and server. What is the reason for using this encoding? (1p)

   (e) Explain the difference between a type 1 hypervisor and a type 2 hypervisor. (1p)

   (f) What did Popek and Goldberg mean with
   i. A sensitive instruction (1p)
   ii. A privileged instruction (1p)

   (g) What was the Popek and Goldberg criteria for an architecture to be *classically virtualizable*? (1p)

   (h) Give one advantage and one disadvantage with paravirtualization compared to traditional virtualization. (2p)

6. (10 p)

   (a) Many operating systems support semaphores. Explain what different types of semaphores that exist and how they work. Give examples of how to use them to protect critical sections. (2p)

   (b) Describe two strategies for *deadlock prevention*. (4p)

   (c) In distributed systems without shared memory support, coordination can be more difficult than among processes running on the same processor. Explain why and give an example (algorithm sketch) of how to achieve mutual exclusion in such a distributed system. (4p)