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

Operating Systems
DIT 400, EDA092

Exam 2011-08-23

Date, Time, Place: Tuesday 23/8 2011, 14:00-18:00, V building

Course Responsible: Arne Dahlberg, Marina Papatriantafilou (Tel: 772 1705, 772 5413)

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) How is RAID 0+1 and RAID 1+0 affected by two erroneous disks? (2p)

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

   (c) Describe the SSTF scheduling methods for optimizing head movement in disk memories. (2p)

   (d) Give two reasons why it is difficult for an operating system to improve the rotational delay when reading from a disk. (2p)

   (e) A file consists of 4 disk blocks. Which disk operations are needed to modify one block in the file if the filesystem uses RAID 5? (2p)

2. (10 p)

   (a) A method to keep track of used blocks in a filesystem is FAT (File Allocation Table). Describe how the FAT method works. (3p)

   (b) Describe two problems with allowing a filesystem to have a general graph structure (cyclic graph). (2p)

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

   (d) Which system call is used in Linux to create threads that execute in the kernel? (1p)

   (e) What happens when we have user-level threads mapped to a single kernel process and one thread issues a read while there are no data to read? What disadvantage does that imply? (2p)

3. (10 p)

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

   (b) When doing the original implementation of NFS, an extra layer called VFS (Virtual File System) was added to the system interface. Why was the VFS layer added? (1p)

   (c) 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)

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

   (e) Explain how paravirtualization works. (1p)

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

   (g) Explain why virtualization complicates the handling of virtual memory and how the problem is usually solved. (3p)

4. (10 p)

   (a) Describe the LRU and the Clock algorithm as solutions to the page replacement problem. (4p)

   (b) What are the advantages of using the Clock algorithm for page replacement as compared with implementing LRU directly? (3p)

   (c) Why are segmentation and paging sometimes combined into one scheme? (3p)

5. (10 p)

   (a) Describe the Earliest-Deadline-First and the Rate-Monotonic methods for scheduling real-time tasks. (4p)

   (b) Show an example of task set whose deadlines can be met using the Earliest-Deadline-First algorithm but they cannot be met using the Rate-Monotonic algorithm. Use space-time diagrams to illustrate it. (3p)

   (c) Explain the priority inversion problem that arises with priority scheduling in combination with synchronization. How can non-blocking synchronization be of help in coping with the problem? (3p)

6. (10 p)

   (a) Design a solution to the mutual exclusion problem for arbitrary number of processes/threads in a multiprocessor system where the following *atomic instruction*, called Fetch-and-Increment is available by the hardware. Discuss carefully the correctness and the other properties of your solution. (5p)

```
int FAI(int *var)
  prev = *var ;
  *var = *var + 1 ;
  return(prev)
```

   (b) Describe a distributed method to solve the mutual exclusion problem in a message passing system based on ordering of events (e.g. the method of Rikard and Agrawala). Discuss carefully the correctness and the other properties of the solution. (5p)