

Review

CSC311 Systems Software

30 November 2015

1. In 1949, some computer scientists already understood the advantages of putting several kinds of devices for storing data and instructions into their machines. In the simplest case, what are the components of a “storage hierarchy?”
2. What is a cache?
3. When and where was the first virtual memory system developed?
4. Denning thinks that the term “virtual memory” probably originated where and when?
5. What is an “Internet edge server?” (You might want to also look up “content delivery network.”)
6. Before the invention of virtual memory, programmers had to work harder. Virtual memory does automatically what programmers once did manually. Explain.
7. In the context of our discussion of virtual memory, what is a “page?”
8. What is “thrashing?”
9. How do the costs of thrashing compare to those of a poor page replacement algorithm?
10. Virtual memory requires the periodic replacement of pages. Ideally, which page should be replaced? (Hint: We should all be fortune tellers!)
11. The IBM System/360 was a hugely successful and influential project. Did the System/360 computers have virtual memory?
12. What was the principal challenge that Denning encountered when he began his efforts to understand and solve the problem of thrashing?
13. Denning suggested that when the level of multiprogramming became too high, thrashing followed. What is the “level of multiprogramming?”
14. Distinguish between temporal clustering and spatial clustering.

15. Let E be the time required in the processor to execute a program, D be the time required to respond to a page fault (that is, the time required to replace a page), and m be the rate of page faults (that is, the number of page faults per unit time).

Then...

- D (a large multiple of the time required to execute one instruction) and E (the product of the number of instructions in the program and the time required to execute each instruction) are fixed
- the design of the virtual memory system will determine the value of m
- mE is the number of page faults that will occur during the execution of the program.
- mED is the time spent responding to page faults
- $E + mED$ is the time required to execute the program

...and this expression models efficiency ...

$$\text{Efficiency}(m) = \frac{E}{E + mED}$$

Draw a plot of $\text{Efficiency}(m)$.

16. What do the following terms mean?

CPU

LRU

MMU

PT

TLB

17. Let x be the amount of memory given to a process and L be rate of page faults for the process. What is the shape of a plot of L as a function of x ?
18. Here is Denning's 1980 definition of the locality principle. Express this idea as clearly and concisely as you can in your own words.
- computational processes pass through a sequence of locality sets
 - the locality sets can be inferred by applying a distance function to a program's address trace observed during a backward window
 - memory management is optimal when it guarantees each program that its locality sets will be present in high speed memory