How to streamline your life (lessons from computer architecture).

COS 116

4/1/2008

Instructor: Sanjeev Arora

Lesson 1: Caching

(and the 80-20 rule)

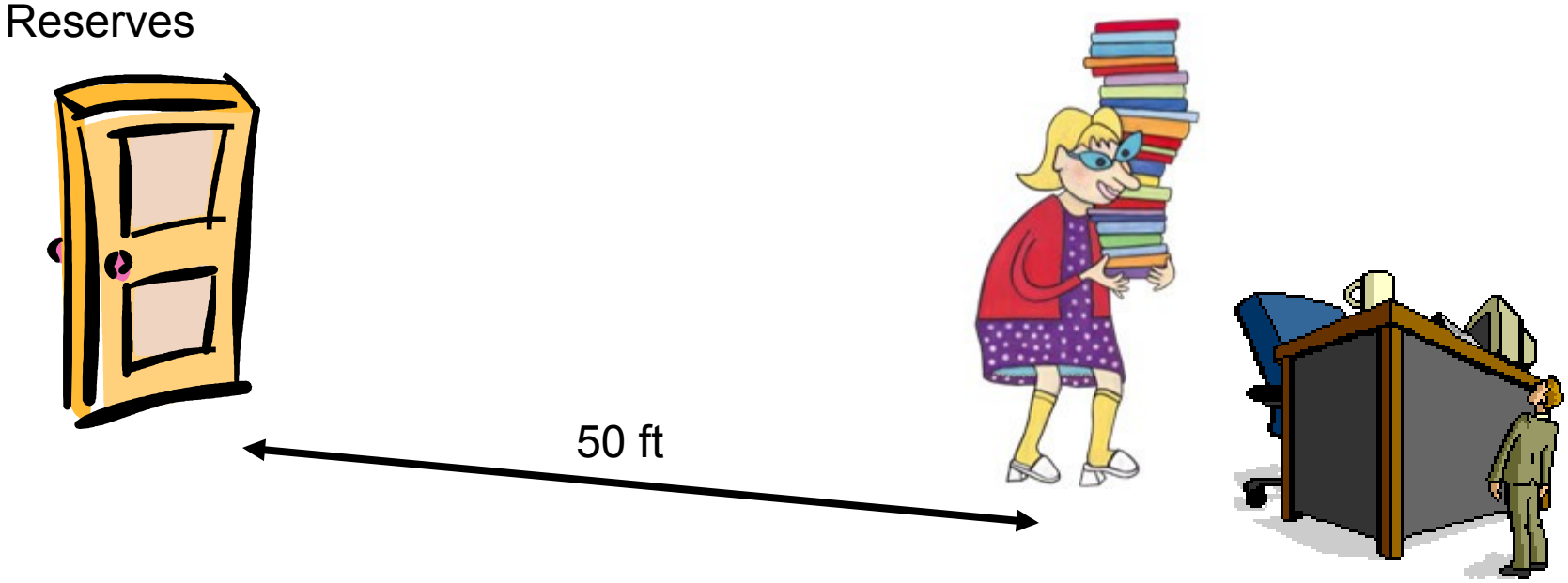
Customer Rating

4.7 out of 5

XPS 420

Intel® Core™2 Q6600 Quad-Core (8MB L2 cache, 2.4GHz, 1066FSB)

The Tired Librarian

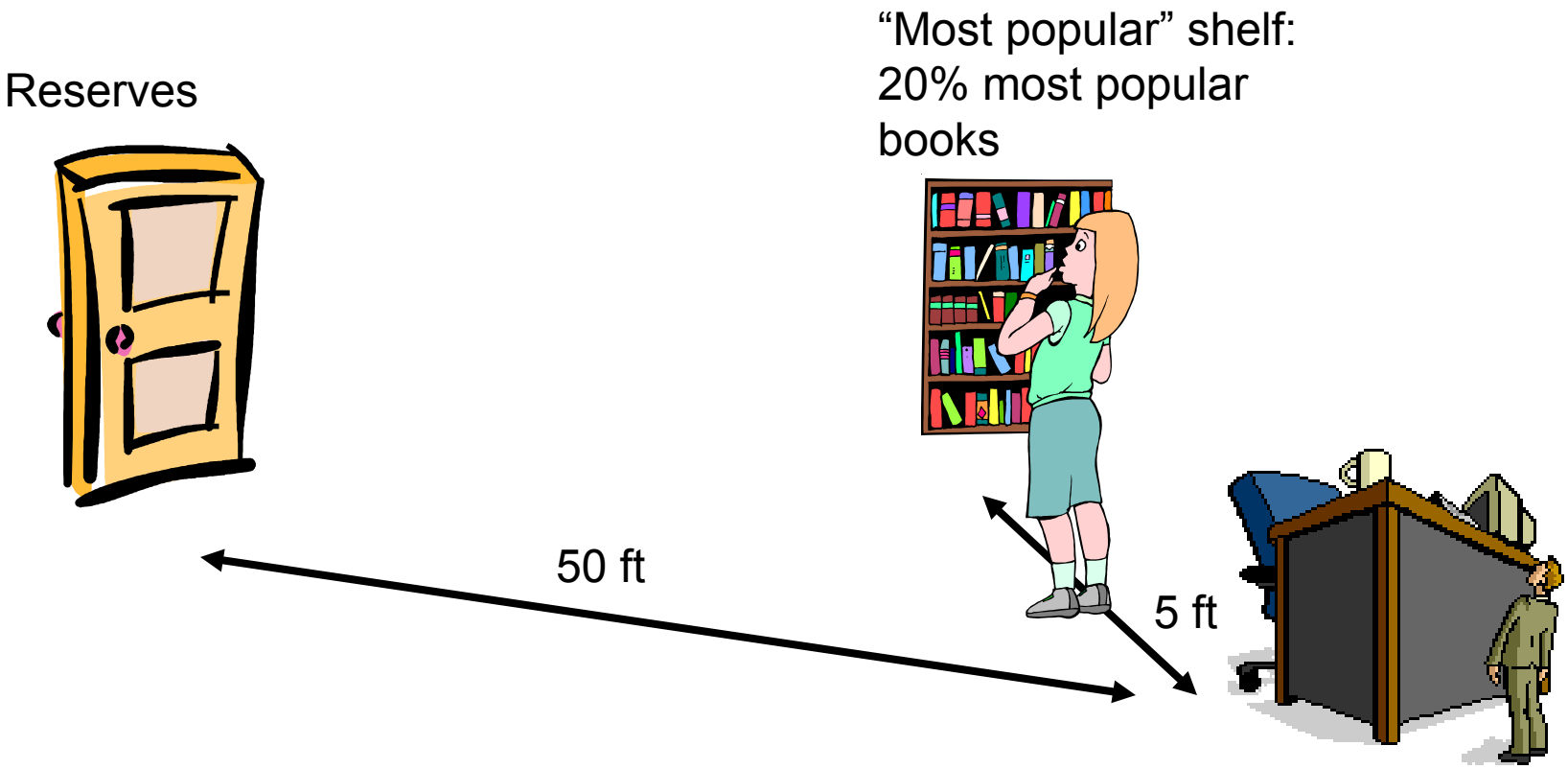


- 1000 checkouts/returns per day
- Distance covered = 50 x 2 x 1000 = 100,000 feet ~ 20 miles
- Please help!!!

80-20 “Rule”

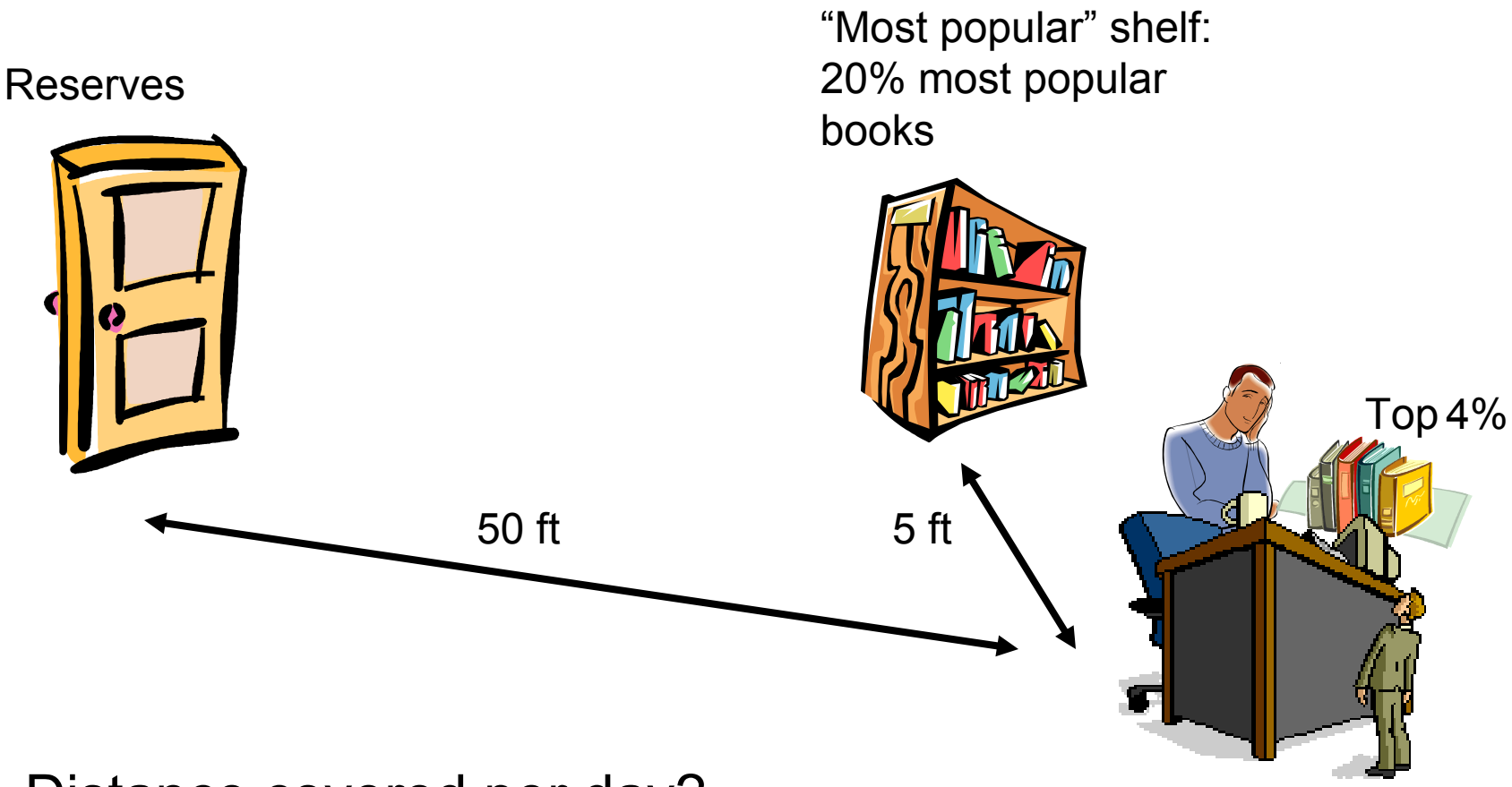
- Pareto [1906]: 20% of the people own 80% of the wealth
- Juran [1930’s]: 20% of the organization does 80% of the work

Better Arrangement



- Distance covered per day?

Even better arrangement?



- Distance covered per day?

Discussion Time

- Is the librarian’s problem solved?

How to predict the 20% most popular books for next day?

- In general, no easy solution
- In practice, use rules of thumb
  - Example: “Least Recently Used”. When you need to create space on the desk (or shelf), move out the book that was used least recently
  - Many others (LRU is computationally expensive)

New and improved

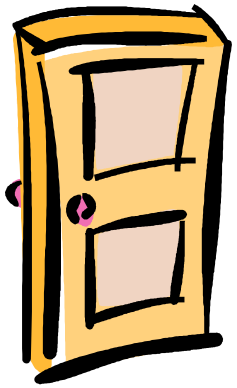
Connection to Computer Organization

- Speed vs cost of various memories

	Cost: $ / GB	Speed: GB/s
Hard drive	0.50	1
RAM	100	5
On-chip memory for CPU (L2 Cache)	40000	15

Computer ~~Librarian~~ arrangement

Reserves Disk



"Most popular" shelf: 20% most popular books Memory

CPU Top 4% Cache

Often, today's computers have even more levels of caching

Moral

- Performance:
  - Speed is close to that of fastest memory (cache)
  - Overall capacity is that of largest memory (disk)

Question

- How does the same program (.exe file) run on different PCs with different memory configurations?

- Answer: “Virtual Memory”
  - All programs live a fiction: allowed to pretend it has 2^{64} bytes of memory
  - Illusion is preserved by hardware

Goodbye Lenin

"The German Democratic Republic lives on – in 79 m²!"

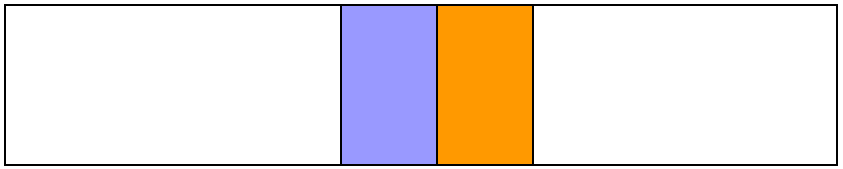
(*Die DDR lebt weiter – auf 79 qm!*)

Virtual Memory

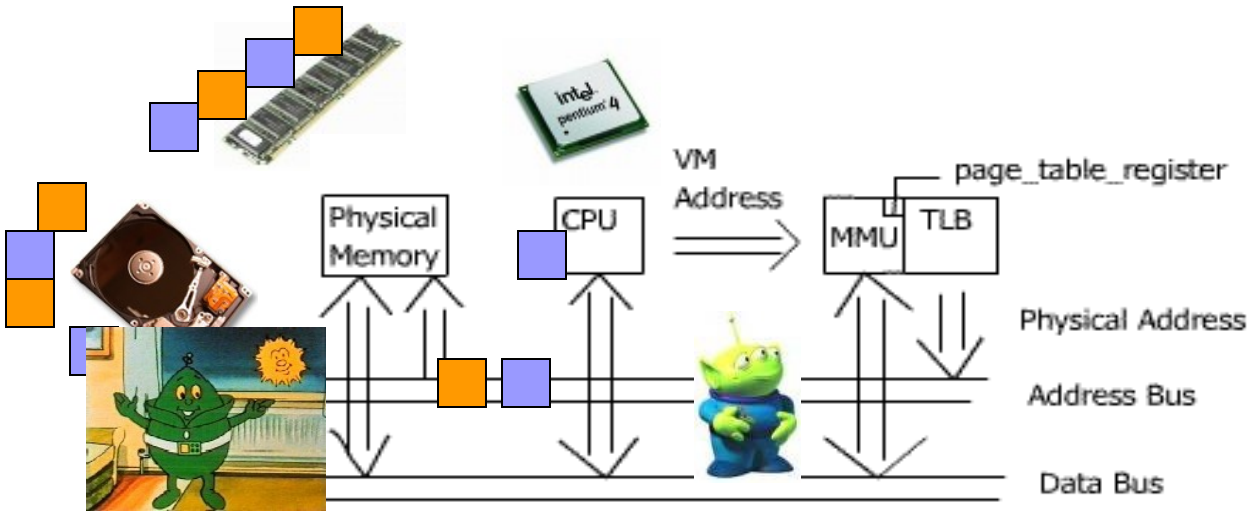
- Program's view:

Powerpoint

Memory:



- Underlying truth:



Lesson 2: Multitasking

- "The Multitasking Generation"

An Evening's Tasks for a Gen-M'er

- Homework
- Listen to music
- Instant Messaging
- Call Mom (goes to bed by 11 PM!)
- Answer phone
- Read a bit more of Joyce's *Ulysses*
- Watch the Daily Show

- How do you do it all?

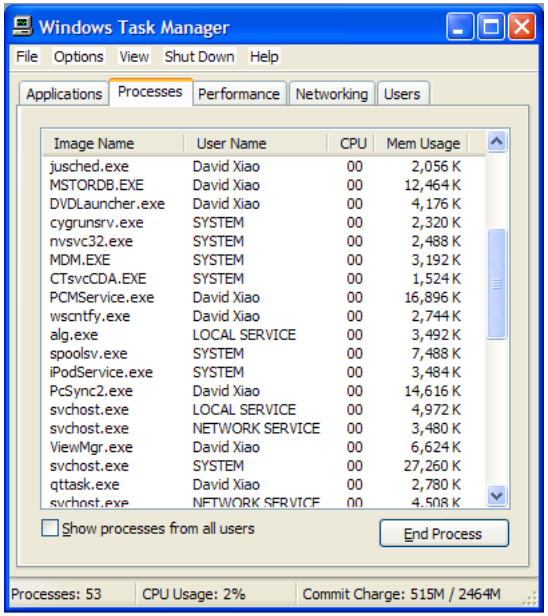
Scheduler's objectives

- Fairness
- Timeliness
- Critical tasks processed promptly
- Low overhead

Discussion Time

How can one achieve these (often conflicting) goals?

Tasks done by my PC last night



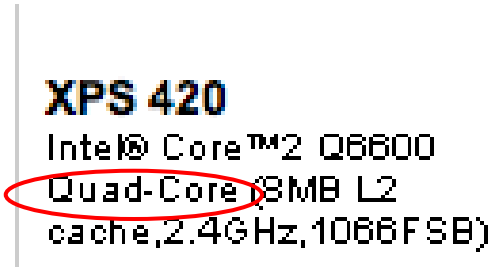
- Word processing
- Play CD
- Download news updates
- Download email
- Run clock
- Hidden tasks: handle network traffic, manage disk and RAM traffic, scheduler, etc.

Managed by “Operating System”
(WinXP, Linux, MacOS, etc.)

Multitasking versus Parallel Processing

Multitasking: A single CPU handles many tasks by switching rapidly among them.
(e.g., all Wintel machines since early 1990s; all Unix machines since the 1970s)

Parallel Processing: Multiple CPUs that do the work of a single CPU. (But, 4 CPUs do not necessarily mean 4x speed.)



The Legal View….

UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

UNITED STATES OF AMERICA,

Plaintiff

v.

MICROSOFT CORPORATION,

Defendant.

Civil Action No. 98-1232 (CKK)

FINAL JUDGMENT
(November 12, 2002)

WHEREAS, plaintiffs United States of America ("United States") and the States of New York, Ohio, Illinois, Kentucky, Louisiana, Maryland, Michigan, North Carolina and Wisconsin and defendant Microsoft Corporation ("Microsoft"), by their respective attorneys, have consented to the entry of this Final Judgment;

One main point studied by the judge:

What is an OS?