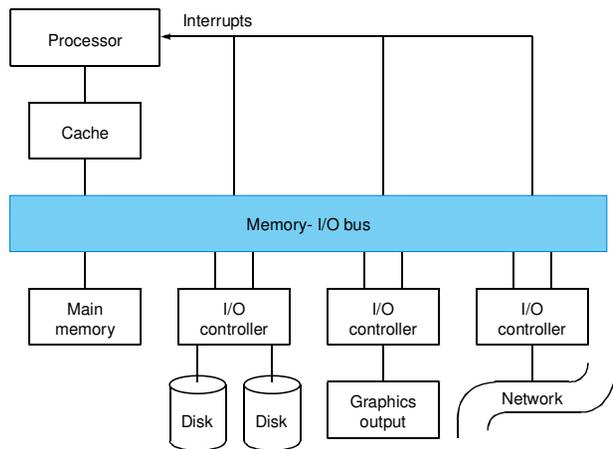

IC220 Set #18: Storage and I/O (Chapter 6)

1

Big Picture



3

ADMIN

- Reading – Chapter 6
 - Including RAID (6.9) but don't stress memorizing the names of the levels
 - Can skip 6.10 and 6.11

2

I/O

- Important but neglected
 - “The difficulties in assessing and designing I/O systems have often relegated I/O to second class status”
 - “courses in every aspect of computing, from programming to computer architecture often ignore I/O or give it scanty coverage”
 - “textbooks leave the subject to near the end, making it easier for students and instructors to skip it!”
- GUILTY!
 - we won't be looking at I/O in much detail
 - be sure and read Chapter 6 carefully
 - Later – IC322: Computer Networks

4

Outline

- A. Overview
- B. Physically connecting I/O devices to Processors and Memory (8.4)
- C. Interfacing I/O devices to Processors and Memory (8.5)
- D. Performance Measures (8.6)
- E. Disk details/RAID (8.2)

5

(A) I/O Overview

- Can characterize devices based on:
 1. behavior
 2. partner (who is at the other end?)
 3. data rate
- Performance factors:
 - access latency
 - throughput
 - connection between devices and the system
 - the memory hierarchy
 - the operating system
- Other issues:
 - Expandability, dependability

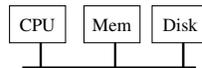
6

(B) Connecting the Processor, Memory, and other Devices

Two general strategies:

1. Bus: _____ communication link

Advantages:



Disadvantages:

2. Point to Point Network: _____ links
Use switches to enable multiple connections

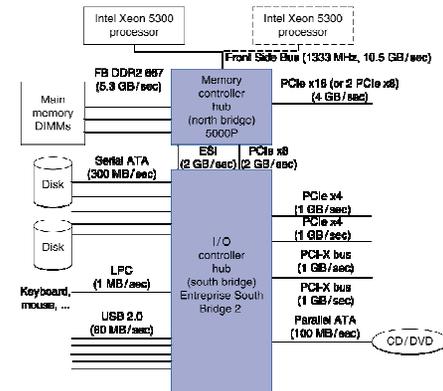
Advantages:



Disadvantages:

7

Typical x86 PC I/O System



8

(B) Bus Basics – Part 1

- Types of buses:
 - Processor-memory
 - Short, high speed, fixed device types
 - custom design
 - I/O
 - lengthy, different devices
 - Standards-based e.g., USB, Firewire
 - Connect to proc-memory bus rather than directly to processor
- Only one pair of devices (sender & receiver) may use bus at a time
 - Bus _____ decides who gets the bus next based on some _____ strategy
 - May incorporate priority, round-robin aspects
- Have two types of signals:
 - “Data” – data or address
 - Control

9

I/O Bus Examples

	Firewire	USB 2.0	PCI Express	Serial ATA	Serial Attached SCSI
Intended use	External	External	Internal	Internal	External
Devices per channel	63	127	1	1	4
Data width	4	2	2/lane	4	4
Peak bandwidth	50MB/s or 100MB/s	0.2MB/s, 1.5MB/s, or 60MB/s	250MB/s/lane 1x, 2x, 4x, 8x, 16x, 32x	300MB/s	300MB/s
Hot pluggable	Yes	Yes	Depends	Yes	Yes
Max length	4.5m	5m	0.5m	1m	8m
Standard	IEEE 1394	USB Implementers Forum	PCI-SIG	SATA-IO	INCITS TC T10

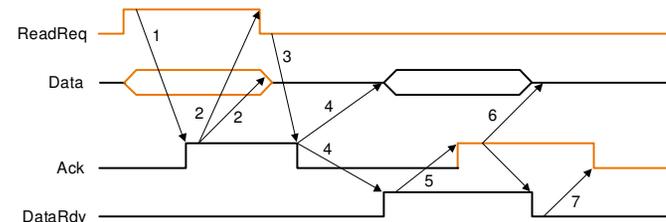
10

(B) Bus Basics – Part 2

- Clocking scheme:
 1. _____
 - + Use a clock, signals change only on clock edge
 - + Fast and small
 - All devices must operate at same rate
 - Requires bus to be short (due to clock skew)
 2. _____
 - + No clock, instead use “handshaking”
 - + Longer buses possible
 - + Accommodate wide range of device
 - more complex control

11

Handshaking example – CPU read from memory



1. CPU requests read
2. Memory acknowledges, CPU deasserts request
3. Memory sees change, deasserts Ack
4. Memory provides data, asserts DataRdy
5. CPU grabs data, asserts Ack
6. Memory sees Ack, deasserts DataRdy
7. CPU sees change, deasserts Ack

12

(C) Processor-to-device Communication

How does CPU send information to a device?

1. Special I/O instructions

x86: `inb / outb`

How to control access to I/O device?

2. Use normal load/instructions to special addresses

Called _____

Load/store put onto bus

Memory ignores them (outside its range)

Address may encode both device ID and a command

How to control access to I/O device?

13

(C) Device-to-processor communication

How does device get data to the processor?

1. CPU periodically checks to see if device is ready: _____

- CPU sends request, keep checking if done
- Or just checks for new info (mouse, network)

2. Device forces action by the processor when needed: _____

- Like an unscheduled procedure call
- Same as "exception" mechanism that handles TLB misses, divide by zero, etc.

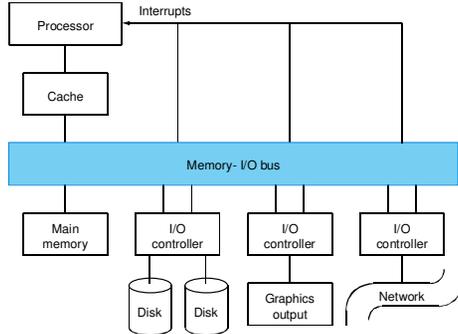
3. DMA:

- Device sends data directly to memory w/o CPU's involvement
- Interrupts CPU when transfer is complete

14

DMA Issues

What could go wrong?



15

(D) I/O's impact on performance

- Total time = CPU time + I/O time
- Suppose our program is 90% CPU time, 10% I/O. If we improve CPU performance by 10x, but leave I/O unchanged, what will the new performance be?
- Old time = 100 seconds
- New time =

16

(D) Measuring I/O Performance

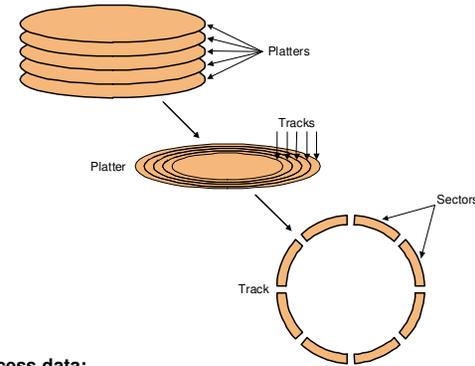
- Latency?
- Throughput?
- Throughput with maximum latency?

- Transaction processing benchmarks
 - TPC-C
 - TPC-H
 - TPC-W

- File system / Web benchmarks
 - “Make” benchmark
 - SPECFS
 - SPECWeb
 - SPECPower

17

(E) Disk Drives



- To access data:
 - seek: position head over the proper track (3 to 14 ms. avg.)
 - rotational latency: wait for desired sector (.5 / RPM)
 - transfer: grab the data (one or more sectors) 30 to 80 MB/sec

18

(E) RAID

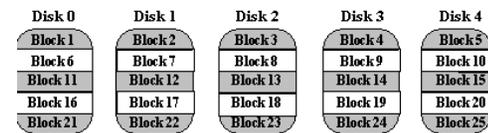
- _____

- Idea: lots of cheap, smaller disks
- Small size and cost makes easier to add redundancy
- Multiple disks increases read/write bandwidth

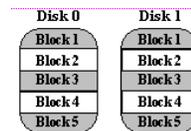
19

RAID

RAID 0 – “striping”, no redundancy



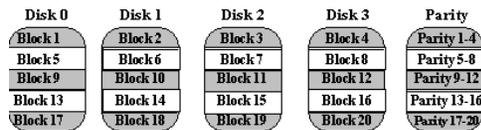
RAID 1 – mirrored



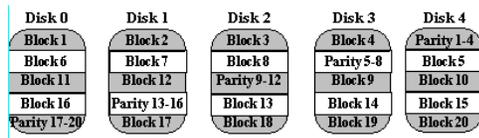
20

RAID

RAID 4 – Block-interleaved parity



RAID 5 – Distributed Block-interleaved Parity



21

Flash Storage

- Nonvolatile semiconductor storage
 - 100x – 1000x faster than disk
 - Smaller, lower power, more robust
 - But more \$/GB (between disk and DRAM)

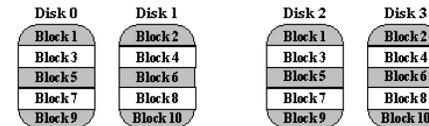


- Flash bits wears out after 1000's of writes
 - Not suitable for direct RAM or disk replacement
 - Wear leveling: remap data to less used blocks
 - Result: “solid-state hard drive”

23

RAID

RAID 10 – Striped mirrors



- Key point – still need to do other backups (e.g. to tape)
 - Provides protection from limited number of disk failures
 - No protection from human failures!

22

Fallacies and Pitfalls

- Fallacy: the rated mean time to failure of disks is 1,200,000 hours, so disks practically never fail.
- Fallacy: magnetic disk storage is on its last legs, will be replaced.
- Fallacy: A GB/sec bus can transfer 1 GB of data in 1 second.
- Pitfall: Moving functions from the CPU to the I/O processor, expecting to improve performance without analysis.

24