
IC220 Set #19: Storage and I/O (Chapter 8)

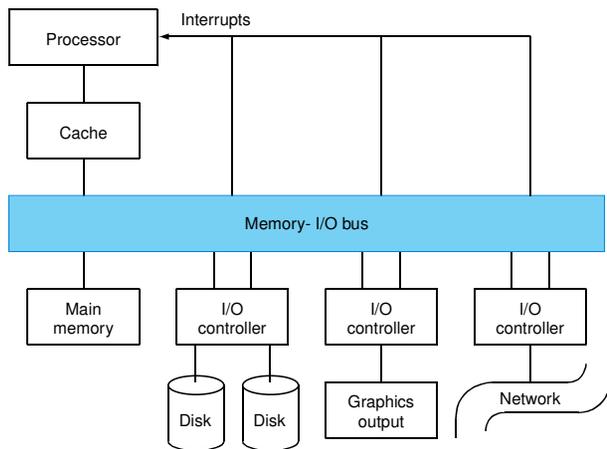
1

ADMIN

- Reading – Chapter 8
 - Including RAID (8.2) but don't stress memorizing the levels
 - Can skip 8.8

2

Big Picture



3

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 8 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

(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

(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) Bus Basics – Part 1

- Types of buses:
 - Process-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 process
- 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

8

I/O Bus Standards

- Today we have two dominant I/O bus standards:

Characteristic	Firewire (1394)	USB 2.0
Bus type	I/O	I/O
Basic data bus width (signals)	4	2
Clocking	asynchronous	asynchronous
Theoretical peak bandwidth	50 MB/sec (Firewire 400) or 100 MB/sec (Firewire 800)	0.2 MB/sec (low speed), 1.5 MB/sec (full speed), or 60 MB/sec (high speed)
Hot pluggable	yes	yes
Maximum number of devices	63	127
Maximum bus length (copper wire)	4.5 meters	5 meters
Standard name	IEEE 1394, 1394b	USE Implementors Forum

FIGURE 8.9 Key characteristics of two dominant I/O bus standards.

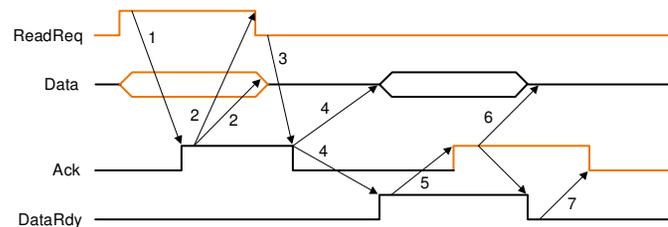
9

(B) Bus Basics – Part 2

- Clocking scheme:
 1. clocked
 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. handshaking
 No clock, instead use “handshaking”
 + Longer buses possible
 + Accommodate wide range of device
 - more complex control

10

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

11

(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/store instructions to special addresses

Called IOPL

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?

12

(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

13

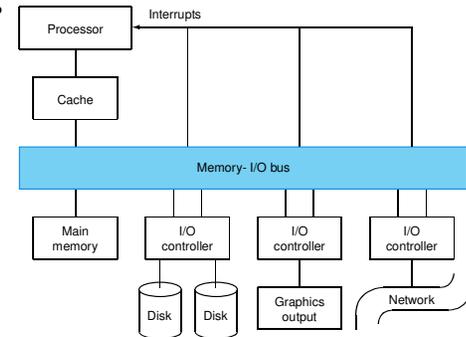
(D) I/O's impact on performance

- 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 =

15

DMA and Cache Coherency

What could go wrong?



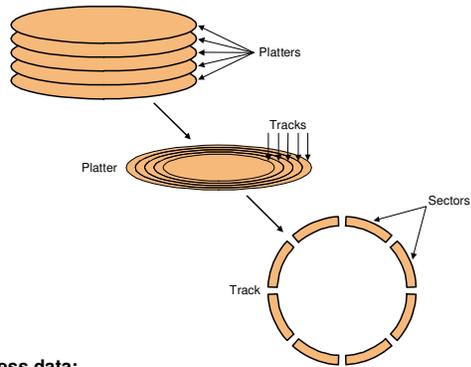
14

(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
 - SPECSFS
 - SPECWeb

16

(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

17

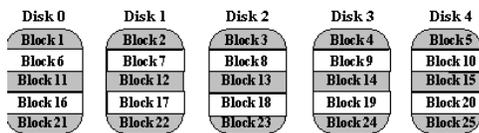
(E) RAID

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

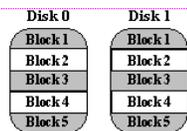
18

RAID

RAID 0 – “striping”, no redundancy



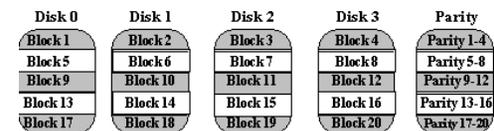
RAID 1 – mirrored



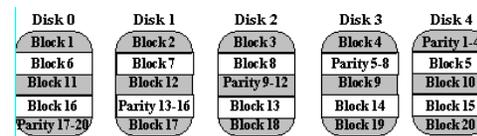
19

RAID

RAID 4 – Block-interleaved parity



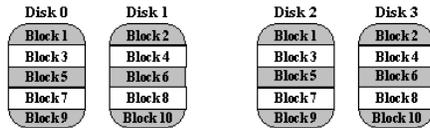
RAID 5 – Distributed Block-interleaved Parity



20

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!

21

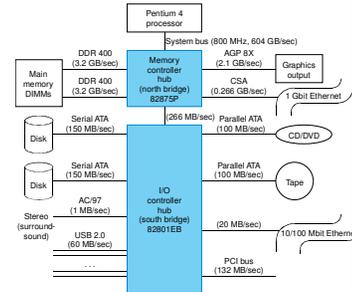
Fallacies and Pitfalls

- **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 100 MB/sec bus can transfer 100 MB/sec.
- **Pitfall:** Moving functions from the CPU to the I/O processor, expecting to improve performance without analysis.

23

Pentium 4

- I/O Options



	875P chip set	845GL chip set
Target segment	Performance PC	Value PC
System bus (64 bit)	800/533 MHz	400 MHz
Memory controller hub ("north bridge")		
Package size, pins	425 × 425 mm, 1100	37.5 × 37.5 mm, 700
Memory speed	DDR 400/533/266 SDRAM	DDR 266/200, PC133 SDRAM
Memory buses, widths	2 × 72	1 × 64
Number of DIMMs, DRAM MDR support	4, 128/256/512 MBs	2, 128/256/512 MBs
Maximum memory capacity	4 GB	2 GB
Memory error correction available?	yes	no
AGP graphics bus, speed	yes, 8x or 4x	no
Graphics controller	external	internal (Extreme Graphics)
CPU digital Ethernet interface	yes	no
South bridge interface, speed (8 bit)	266 MHz	200 MHz
I/O controller hub ("south bridge")		
Package size, pins	31 × 31 mm, 450	31 × 31 mm, 421
PCI bus: width, speed, masters	32 bit, 33 MHz, 6 masters	32 bit, 33 MHz, 6 masters
Ethernet MAC controller, interface	100/10 Mb/s	100/10 Mb/s
USB 2.0 ports, controllers	6, 4	6, 3
ATA 100 ports	2	2
Serial ATA 150 controller, ports	yes, 2	no
RAID 0 controller	yes	no
AC/97 audio controller, interface	yes	yes
I/O management	SMbus 2.0, GPIO	SMbus 2.0, GPIO

FIGURE 4.12 Two Pentium 4 I/O chip sets from Intel. The 845GL north bridge uses many fewer pins than the 875P by having just one memory bus and by omitting the AGP bus and the CPU digital Ethernet interface. Note that the serial nature of USB and Serial ATA means that two more USB ports and two more Serial ATA ports need just 39 more pins in the south bridge of the 875P versus the 845GL chip sets.

22