

COMPUTER ARCHITECTURE - CHAPTER 7 AND 8

Timothy Hnat

Department of Computer Science
University of Virginia

Oct 18, 2007



INTRODUCTION

- Does I/O Performance Matter?
- Does CPU Performance Matter?



INTRODUCTION

- Does I/O Performance Matter?
- Does CPU Performance Matter?

ANSWER

- Throughput - number of tasks completed per hour
- Interactive processing
- Dependability vs performance.



TYPES OF STORAGE

- Magnetic Disks
- Optical Disks
- Magnetic Tapes
- Automated Tape Libraries
- Flash Memory



TYPES OF STORAGE

- Magnetic Disks
- Optical Disks
- Magnetic Tapes
- Automated Tape Libraries
- Flash Memory

ANSWER

- Magnetic Disks - Long-term, nonvolatile storage
- Tapes - Long time for access (Sequential)
- Flash memory - Compare this to disks



BUSES - CONNECT I/O DEVICES TO CPU/MEMORY

- What are the advantages and disadvantages of buses?
- Why not apply Moore's law to buses?



BUSES - CONNECT I/O DEVICES TO CPU/MEMORY

- What are the advantages and disadvantages of buses?
- Why not apply Moore's law to buses?

ANSWER

- Disadvantage - Communication bottleneck
- Buses are limited by physical constraints
- Table on page 694



BUSES CONTINUED

- What are the difference between synchronous and asynchronous buses?
- Drawbacks for both?
- For a given bus speed, which is faster?



BUSES CONTINUED

- What are the difference between synchronous and asynchronous buses?
- Drawbacks for both?
- For a given bus speed, which is faster?

ANSWER

- Synchronous buses contain a clock in the control lines and a fixed protocol
- Synchronous buses are fast and inexpensive
- Drawback: Clock skew and single bus speed
- Asynchronous buses contain no clock and allow variable protocols
- Handshake protocol for two way communications



INTERFACING STORAGE DEVICES TO THE CPU

- What are the three ways to interface to the CPU?



INTERFACING STORAGE DEVICES TO THE CPU

- What are the three ways to interface to the CPU?

ANSWER

- Memory-mapped I/O
- Polling
- Interrupt-driven I/O



DEFINING FAILURE

- Define dependability, failure, latent errors



DEFINING FAILURE

- Define dependability, failure, latent errors

ANSWER

- Dependability is the quality of delivered service such that reliance can justifiably be placed on this service.
- Failures occur because of defects in modules
- Latent errors occur when it is activated
- A fault can create more than one latent error



TYPES OF FAULTS

- What are the three types of faults?



TYPES OF FAULTS

- What are the three types of faults?

ANSWER

- Transient - exist for limited time and not recurring
- Intermittent - cause a system to oscillate between faulty and fault-free operation
- Permanent - do not correct themselves with passing time



RELIABILITY

- What are the four methods for improving reliability?



RELIABILITY

- What are the four methods for improving reliability?

ANSWER

- Fault avoidance - prevents fault occurrence
- Fault tolerance - provides redundancy
- Error removal - minimize latent errors
- Error forecasting - estimate presence of faults



THROUGHPUT VERSUS RESPONSE TIME

- What is the difference?



THROUGHPUT VERSUS RESPONSE TIME

- What is the difference?

ANSWER

- Throughput is the amount of information you can process in a given time
- Response Time is the amount of time waiting in a buffer



QUEUING THEORY

- What is M/M/1?



QUEUING THEORY

- What is M/M/1?

ANSWER

- M = Exponentially random request arrival
- M = exponentially random request arrival
- 1 = single server



BENCHMARKS

- What is OLTP?
- Where is the bottleneck?



BENCHMARKS

- What is OLTP?
- Where is the bottleneck?

ANSWER

- OLTP - Online Transaction Processing
- I/O Rate - the number of disk accesses per second



DMA

- What are the problems with DMA and Virtual Memory?
- How do you solve them?



DMA

- What are the problems with DMA and Virtual Memory?
- How do you solve them?

ANSWER

- Transferring a buffer that is larger than one page will cause problems
- The operating system removes some of the pages from memory
- Solution: DMA pins pages in memory for the duration of I/O
- Solution: All DMA to write to virtual addresses



EXAMPLES

- Pages 742-754



FALLACIES AND PITFALLS

- The rated mean time to failure of disks is 1,20,000 hours or almost 140 years, so disks practically never fail.



FALLACIES AND PITFALLS

- The rated mean time to failure of disks is 1,20,000 hours or almost 140 years, so disks practically never fail.

ANSWER

- MTTF is computed by running disks for a few months and computing the total time divided by the number that failed
- The percentage of disks that fail would be more useful.
- The expected lifetime of a disk is 5 years or 43,800 hours



FALLACIES AND PITFALLS

- Components fail fast



FALLACIES AND PITFALLS

- Components fail fast

ANSWER

- Most components will develop transient errors before they fail completely.



FALLACIES AND PITFALLS

- Computer systems achieve 99.999



FALLACIES AND PITFALLS

- Computer systems achieve 99.999

ANSWER

- Only in special tests
- Does not include scheduled downtime



FALLACIES AND PITFALLS

- Where a function is implemented affects its reliability



FALLACIES AND PITFALLS

- Where a function is implemented affects its reliability

ANSWER

- In theory RAID works fine in software
- Hardware puts more emphasis on testing correctness



FALLACIES AND PITFALLS

- Semiconductor memory will soon replace magnetic disks in desktop and server computer systems



FALLACIES AND PITFALLS

- Semiconductor memory will soon replace magnetic disks in desktop and server computer systems

ANSWER

- Disks are growing even faster than DRAM



FALLACIES AND PITFALLS

- Since head-disk assemblies of disks are the same technology independent of the disk interface, the disk interface matters little in price.



FALLACIES AND PITFALLS

- Since head-disk assemblies of disks are the same technology independent of the disk interface, the disk interface matters little in price.

ANSWER

- SCSI vs IDE is not much of a difference now



FALLACIES AND PITFALLS

- The time of an average seek of a disk in a computer system is the time for a seek of one-third the number of cylinders.



FALLACIES AND PITFALLS

- The time of an average seek of a disk in a computer system is the time for a seek of one-third the number of cylinders.

ANSWER

- One third is the average seek distance between two random locations.
- Seek time is not linear
- Assumed no locality, data is usually grouped



EXAMPLES IN BOOK

- See the book

