Outline

- Chapter 13: I/O Systems
- Chapter 14: Disk Scheduling
- RAID paper

I/O Systems

- Hardware
  - Bus and interconnects & Controller
  - Host adapter
  - Registers and Memory mapped I/O
  - Polling, Interrupt driven with an interrupt controller
  - Maskable and non-maskable interrupts and priorities
  - Direct Memory Access (DMA)
- Application Interface
  - Character or Block: terminal or disk
  - Sequential or Random access: modem or CD-ROM
  - Synchronous or asynchronous: tape or keyboard
  - Shared or dedicated: tape or keyboard
  - RW, R and W only: disk, CDROM, graphics controller

Kernel I/O

- I/O Scheduling
  - Buffering; double-buffering
    - Copy semantics for async. I/O
  - Caching
  - Spooling and device reservation
  - Error handling
- Performance
  - Front end processors to off load processing

Disk scheduling

- Schedule disk accesses to gain performance
  - FCFS - first come first service
  - SSTF - shorted seek time first
    - starvation
  - SCAN
    - Elevator algorithm
  - CSCAN
    - Restarts from the beginning after each cycle
  - LOOK
    - Look till end of direction
  - CLOOK
- Disk scheduling harder with smart disks that can rearrange bad sectors

RAID

- Reliability vs redundancy
- Performance via parallelism

- Raid 0: striping w/o redundancy
  - No redundancy
  - Good I/O performance
- Raid 1: Mirrored disks
  - Highly redundant
  - Twice read rate, same write performance
- Raid 2: Hamming code ECC
  - Separate disks for data and error correction code
  - Commercially not viable

Raid levels

- Raid 3: bit-interleaved parity organization
  - Data with separate parity disks
- Raid 4: block-interleaved parity
  - Separate parity disk
- Raid 5: Block-interleaved distributed parity
  - Parity data is distributed across all disks
  - Complex implementation on the controller
- Raid 6, 6, 10, 50, 0+1, 53, …

Disk attachment

- Host-attached storage
  - SCSI, Fibre-Channel
- Network attached storage (NAS)
  - Device implements a complete file system
- Storage-Area Networks
  - High speed interconnect
  - Can dynamically reassign disks to other servers
- iSCSI
  - SCSI using IP protocols
  - Security, congestion etc. are issues
- Direct Access File System (DAFS)
  - Emerging standard leveraging Remote Direct Memory Access infrastructure
  - http://www.dafscollaborative.org/