Device Management

INFO 2603
Platform Technologies 1

Week 5: 1-Oct-2018

Outline

- Device Management
  - Features of dedicated, shared, and virtual devices
  - Concepts of blocking and buffering, and how they improve I/O performance
  - Differences in access times in several types of devices
  - Levels of RAID and what sets each apart

Types of Devices

- Three categories: dedicated, shared, and virtual
- Dedicated device
  - Assigned to one job at a time
    - For entire time that job is active (or until released)
    - Examples: tape drives, printers, and plotters
  - Disadvantage
    - Must be allocated for duration of job's execution
    - Inefficient if device is not used 100 percent of time

Types of Devices (cont'd.)

- Shared device
  - Assigned to several processes
    - Example: direct access storage device (DASD)
      - Processes share DASD simultaneously
      - Requests interleaved
  - Device manager supervision
    - Controls interleaving
      - Predetermined policies determine conflict resolution

Types of Devices (cont'd.)

- Virtual device
  - Dedicated and shared device combination
  - Dedicated device transformed into shared device
    - Example: printer
      - Converted by spooling program
      - Spooling: speeds up slow dedicated I/O devices
  - Universal serial bus (USB) controller
    - Interface between operating system, device drivers, applications, and devices attached via USB host
    - Assigns bandwidth to each device: priority-based
      - High, medium, or low priority

Management of I/O Requests

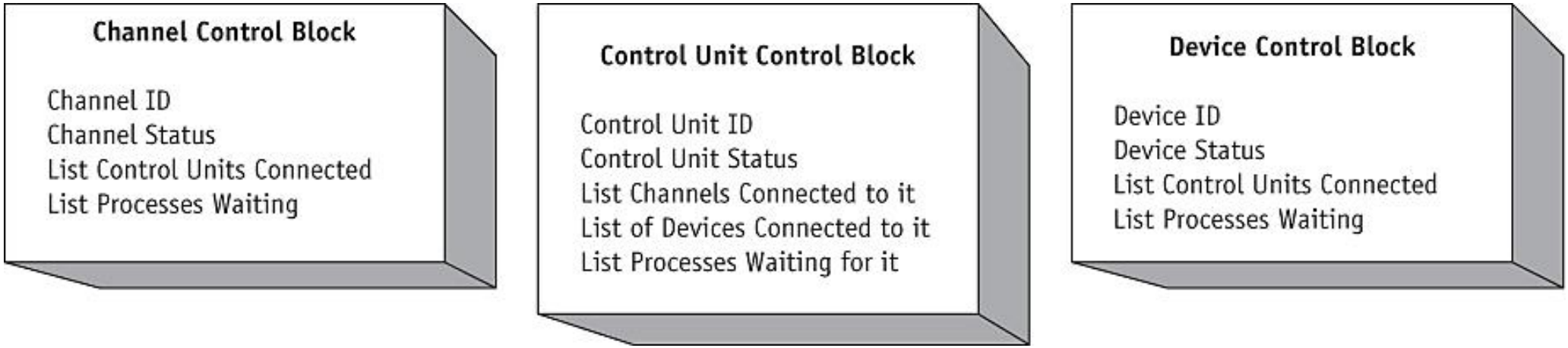
- I/O traffic controller
  - Watches status of devices, control units, channels
  - Three main tasks
    - Determine if path available
    - If more than one path available, determine which one to select
    - If paths all busy, determine when one is available
  - Maintains database containing each unit's status and connections

Management of I/O Requests (cont'd.)

- I/O scheduler
  - Same job as process scheduler (Chapter 4)
  - Allocates devices, control units, and channels
  - If requests greater than available paths
    - Decides which request to satisfy first: based on different criteria
  - In many systems
    - I/O requests not preempted
  - For some systems
    - Allow preemption with I/O request subdivided
    - Allow preferential treatment for high-priority requests

Management of I/O Requests (cont'd.)

- I/O device handler
  - Performs actual data transfer
    - Processes device interrupts
    - Handles error conditions
    - Provides detailed scheduling algorithms
  - Device dependent
  - Each I/O device type has its own device handler algorithm



(figure 7.2)
Each control block contains the information it needs to manage the channels, control units, and devices in the I/O subsystem.

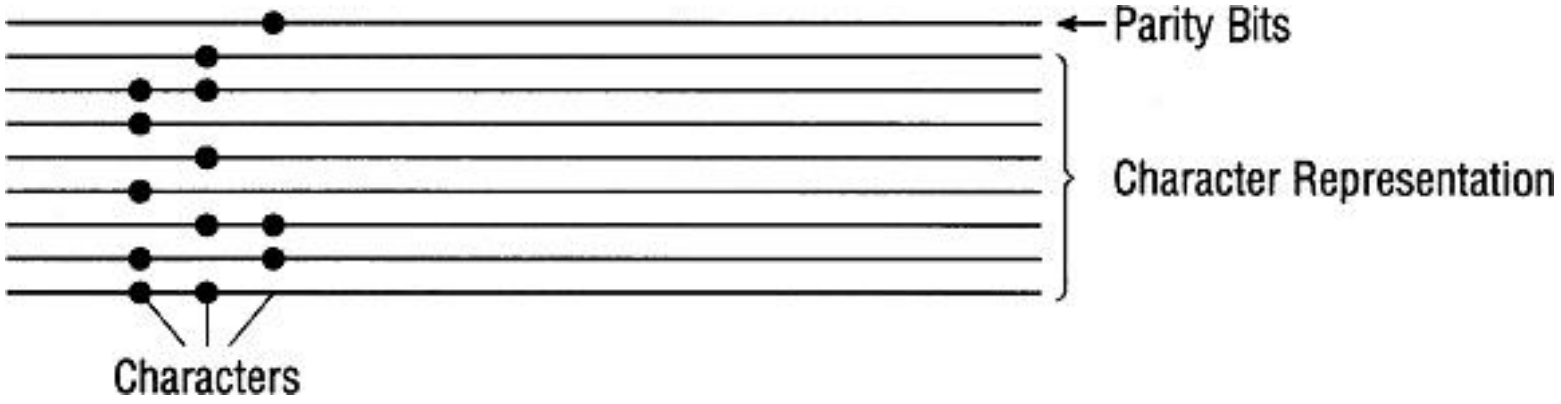
I/O Devices in the Cloud

- Local operating system's role in accessing remote I/O devices
  - Essentially the same role performed accessing local devices
- Cloud provides access to many more devices

Sequential Access Storage Media

- Magnetic tape
  - Early computer systems: routine secondary storage
  - Records stored serially
    - Record length determined by application program
    - Record identified by position on tape
    - Record access
      - Tape rotates passing under read/write head: only when access requested for read or write
    - Time-consuming process

Sequential Access Storage Media (cont'd.)



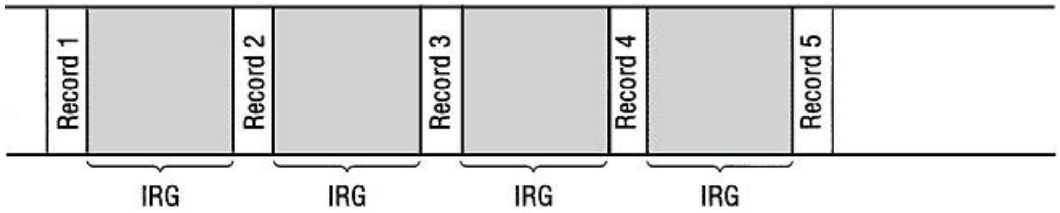
(figure 7.3)
Nine-track magnetic tape with three characters recorded using odd parity. A 1/2-inch wide reel of tape, typically used to back up a mainframe computer, can store thousands of characters, or bytes, per inch.

- Tape density: characters recorded per inch
  - Depends upon storage method (individual or blocked records)

Sequential Access Storage Media (cont'd.)

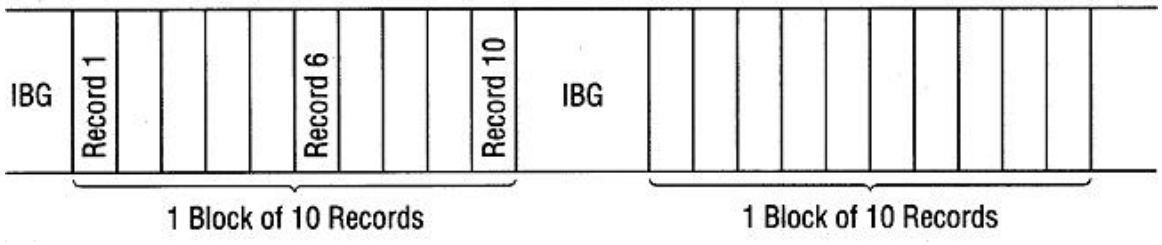
- Interrecord gap (IRG)
  - ½ inch gap inserted between each record
  - Same size regardless of sizes of records it separates
- Blocking: group records into blocks
- Transfer rate: (tape density) x (transport speed)
- Interblock gap (IBG)
  - ½ inch gap inserted between each block
  - More efficient than individual records and IRG
- Optimal block size
  - Entire block fits in buffer

Sequential Access Storage Media (cont'd.)



(figure 7.4)
IRGs in magnetic tape. Each record requires only 1/10 inch of tape. When 10 records are stored individually on magnetic tape, they are separated by IRGs, which adds up to 4.5 inches of tape. This totals 5.5 inches of tape.

(figure 7.5)
Two blocks of records stored on magnetic tape, each preceded by an IBG of 1/2 inch. Each block holds 10 records, each of which is still 1/10 inch. The block, however, is 1 inch, for a total of 1.5 inches.



Sequential Access Storage Media (cont'd.)

- Blocking advantages
  - Fewer I/O operations needed
  - Less wasted tape space
- Blocking disadvantages
  - Overhead and software routines needed for blocking, deblocking, and record keeping
  - Buffer space wasted
    - When only one logical record needed

Sequential Access Storage Media (cont'd.)

- Access time
  - Poor for routine secondary storage except files with very high (90 to 100 percent) sequential activity

Benchmarks	Access Time
Maximum access	2.5 minutes
Average access	1.25 minutes
Sequential access	3 milliseconds

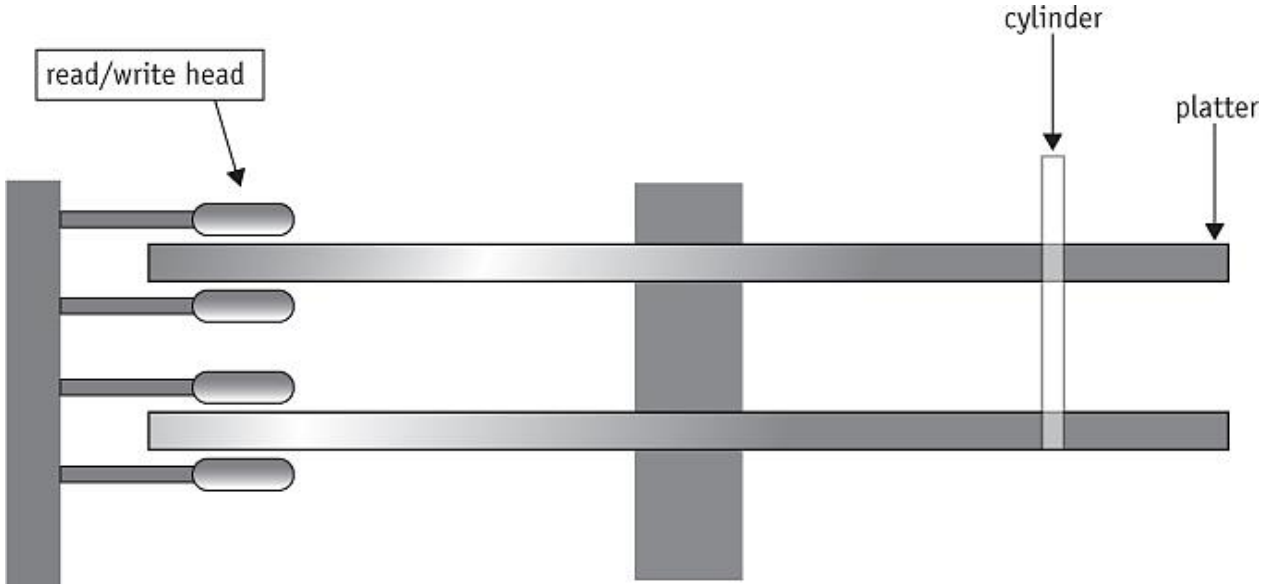
(table 7.1)
Access times for 2400-foot magnetic tape with a tape transport speed of 200 ips.

Direct Access Storage Devices

- Directly read or write to specific disk area
  - Random access storage devices
- Three categories
  - Magnetic disks
  - Optical discs
  - Solid state (flash) memory
- Access time variance
  - Not as wide as magnetic tape
  - Record location directly affects access time

Magnetic Disk Storage

- Computer hard drives
- Single platter or stack of magnetic platters



(figure 7.6)
A disk pack is a stack of magnetic platters. The read/write heads move between each pair of surfaces, and all of the heads are moved in unison by the arm.

Magnetic Disk Storage (cont'd.)

- Two recording surfaces (top and bottom)
- Each surface formatted
  - Concentric tracks: numbered from track 0 on outside to highest track number in center
- Read/write heads move in unison: virtual cylinder
- Accessing a record: system needs three things
  - Cylinder number
  - Surface number
  - Sector number

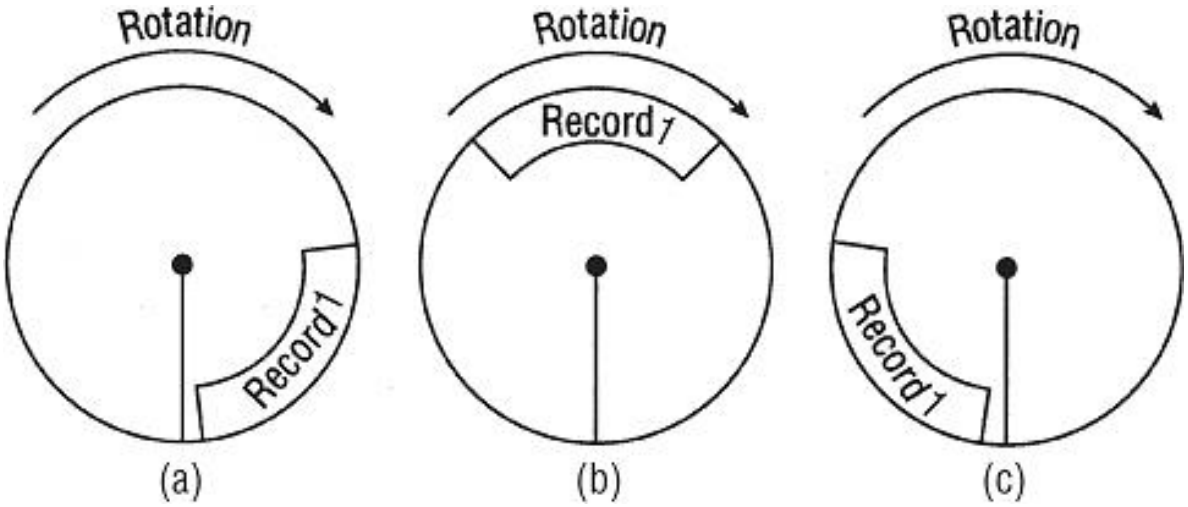
Access Times

- File access time factors
  - Seek time (slowest)
    - Time to position read/write head on track
    - Does not apply to fixed read/write head devices
  - Search time
    - Rotational delay
    - Time to rotate DASD
    - Rotate until desired record under read/write head
  - Transfer time (fastest)
    - Time to transfer data
    - Secondary storage to main memory transfer

Fixed-Head Magnetic Drives

- Record access requires two items
  - Track number and record number
- Total access time = search time + transfer time
- DASDs rotate continuously
  - Three basic positions for requested record
    - In relation to read/write head position
- DASD has little access variance
  - Good candidates: low activity files, random access
- Blocking minimizes access time

(figure 7.8)
As a disk rotates, Record 1 may be near the read/write head and ready to be scanned, as seen in (a); in the farthest position just past the head, (c); or somewhere in between, as in the average case, (b).



Benchmarks	Access Time
Maximum access time	16.8 ms + 0.00094 ms/byte
Average access time	8.4 ms + 0.00094 ms/byte
Sequential access time	Depends on the length of the record (known as the transfer rate)

(table 7.2)
Access times for a fixed-head disk drive at 16.8 ms/revolution.

Movable-Head Magnetic Drives

- Access time = seek time + search time
  + transfer time
- Search time and transfer time calculation
  - Same as fixed-head DASD
- Blocking: good way to minimize access time

Benchmarks	Access Time
Maximum access time	10 ms + 16.8 ms + 0.00094 ms/byte
Average access time	5 ms + 8.4 ms + 0.00094 ms/byte
Sequential access time	Depends on the length of the record, but generally less than 1 ms

(table 7.3)
Typical access times for a movable-head drive, such as a typical hard drive.

Reading Resources

Understanding Operating Systems: Ann McIver McHoes, Ida M. Flynn. 8th Edition. 2017: Chapter 7

Articles:

https://spectrum.ieee.org/view-from-the-valley/computing/hardware/david-patterson-says-its-time-for-new-computer-architectures-and-software-languages