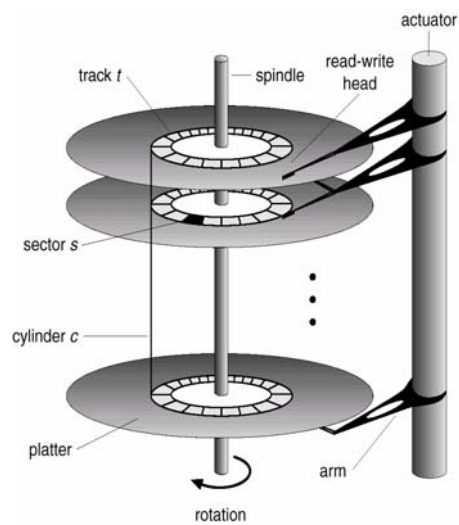


Chapter 14: Mass-Storage

- Overview
- Disk Structure
- Disk Scheduling
- Disk Management
- Tertiary Storage Devices
- Disk Speed, Reliability, and Cost

Overview: Moving-Head Disk Mechanism



Disk Structure

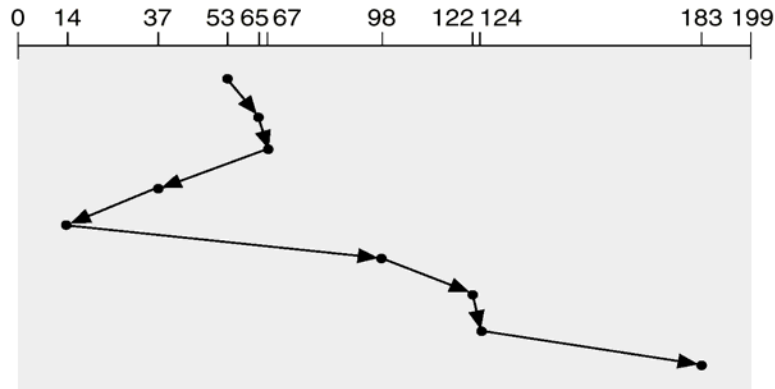
- Disk drives are addressed as large 1-dimensional arrays of *logical blocks*, where the logical block is the smallest unit of transfer.
- The 1-dimensional array of logical blocks is mapped into the sectors of the disk sequentially.
 - Sector 0 is the first sector of the first track on the outermost cylinder.
 - Mapping proceeds in order through that track, then the rest of the tracks in that cylinder, and then through the rest of the cylinders from outermost to innermost.

Disk Scheduling

- The operating system is responsible for using hardware efficiently — for the disk drives, this means having a fast access time and disk bandwidth.
- Access time has two major components
 - *Seek time* is the time for the disk to move the heads (disk arm) to the cylinder containing the desired sector.
 - *Rotational latency* is the additional time waiting for the disk to rotate the desired sector to the disk head.
- Minimize seek time
- Seek time \approx seek distance
- Disk bandwidth is the total number of bytes transferred, divided by the total time between the first request for service and the completion of the last transfer.

SSTF (Cont.)

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53

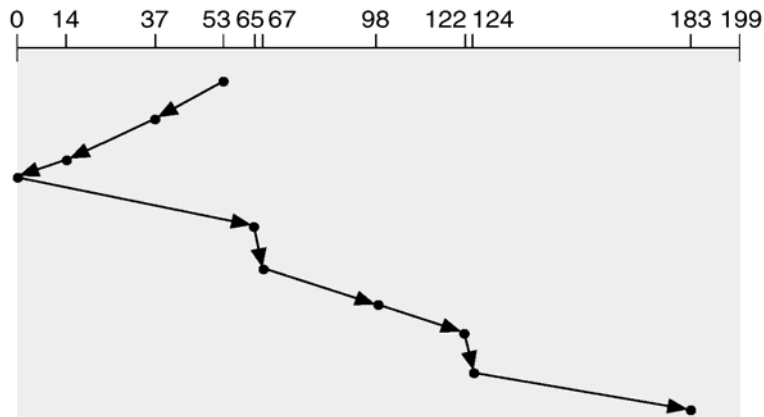


SCAN Scheduling

- The disk arm starts at one end of the disk, and moves toward the other end, servicing requests as it reaches each cylinder until it gets to the other end of the disk, where the head movement is reversed and servicing continues.
- Sometimes called the *elevator algorithm*.
- Scans disk back and forth across the disk.
- Illustration (next example) shows total head movement of 208 cylinders.
- If a request arrives in the queue just in front of the head, it will be served.
- A request arriving just behind the head will have to wait until the arm moves to the end of the disk, reverses direction, and comes back.

SCAN (Cont.)

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53



Operating System Concepts

14.11

C-SCAN (Circular SCAN) Scheduling

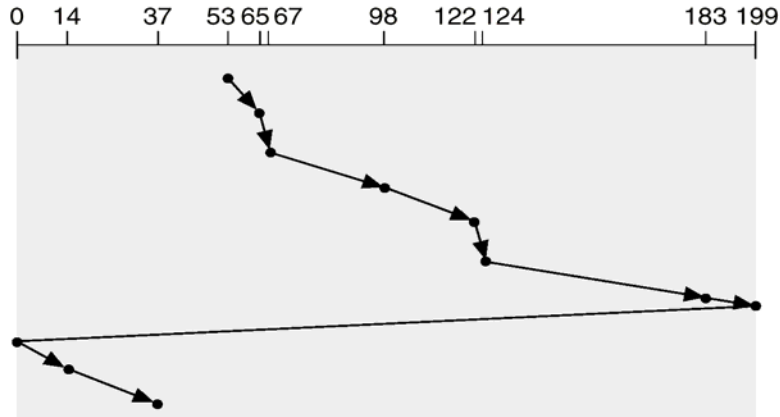
- Provides a more uniform wait time than SCAN.
- Like SCAN, in C-SCAN the head moves from one end of the disk to the other, servicing requests as it goes. But when it reaches the other end, however, it immediately returns to the beginning of the disk, without servicing any requests on the return trip.

Operating System Concepts

14.12

C-SCAN (Cont.)

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53



Operating System Concepts

14.13

LOOK and C-LOOK Scheduling

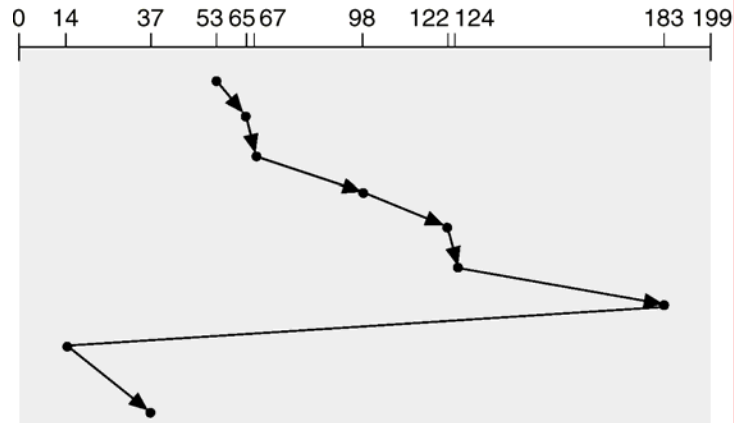
- In both SCAN and C-SCAN the arm of the disk goes all the way to the end of disk (0 and 199).
- LOOK and C-LOOK same as SCAN and C-SCAN respectively, but the arm of the disk only goes as far as the last request in each direction, then reverses direction immediately, without first going all the way to the end of the disk.

Operating System Concepts

14.14

C-LOOK (Cont.)

queue = 98, 183, 37, 122, 14, 124, 65, 67
head starts at 53



Selecting a Disk-Scheduling Algorithm

- SSTF is common.
- SCAN and C-SCAN perform better for systems that place a heavy load on the disk.
- Performance depends on the number and types of requests.
- Requests for disk service can be influenced by the file-allocation method.
- The disk-scheduling algorithm should be written as a separate module of the operating system, allowing it to be replaced with a different algorithm if necessary.
- Either SSTF or LOOK is a reasonable choice for the default algorithm.

Disk Management

- Disk Formatting
- Boot Block
- Bad Blocks

Disk Formatting

- A new magnetic disk is a blank slate: It is just platters of a magnetic recording material.
- *Low-level formatting, or physical formatting* — Dividing a disk into sectors that the disk controller can read and write.
- So, the low-level formatting fills the disk with a special data structure for each sector.
- The data structure for a sector consists of a header, a data area (512 bytes in size), and a trailer.
- The header and trailer contain information used by the disk controller, such as a sector number and an error-correcting code (ECC). The ECC is used to identify corrupted sectors.

Disk Formatting (Cont.)

- Most hard disks are low-level formatted at the factory as a part of the manufacturing process to test the disk.
- To use a disk to hold files, the operating system still needs to record its own data structures on the disk. It does so in two steps:
 - Partition the disk into one or more groups of cylinders to for example hold the operating system's executable code in one partition and user files in another partition.
 - Logical formatting or "making a file system": The O.S. stores the initial file-system data structures onto the disk. These data structures may include the maps of free and allocated space.

Boot Block

- For a computer to start running (when it is powered up or rebooted) it needs to have an initial program to run. This program called bootstrap program.
- The bootstrap program initializes the system, from CPU registers to device controllers and the contents of main memory, and then starts the operating system.
- The bootstrap is stored in read-only memory (ROM):
 - ROM needs no initialization.
 - ROM is at a fixed location that the processor can start executing when powered up or reset.
 - ROM cannot be infected by a computer virus.

Boot Block (Cont.)

- Problem: Changing bootstrap code requires changing the ROM hardware chips. For this reason, most systems store a tiny bootstrap loader program in the boot ROM, whose only job is to bring in a full bootstrap program from disk.
- The full bootstrap program can be changed easily: A new version is simply written onto the disk.
- The full bootstrap program is stored in a partition called the boot blocks, at a fixed location on the disk.
- A disk that has a boot partition is called a boot disk or system disk.

Bad Blocks

- Sometimes disks come from factory with bad blocks. These bad blocks are handled in variety of ways.
- Example: The MS-DOS format command does a logical format, and scans the disk to find bad blocks.
 - If format finds a bad block, it writes a special value into the corresponding FAT (File-Allocation Table is a method of disk-space allocation) entry to tell the allocation routines not to use that block.
 - If blocks go bad during normal operation, a special program (chkdsk) must be run manually to search for the bad blocks and to lock them away as before.
- Data that resided on the bad blocks usually are lost.

Bad Blocks (Cont.)

- A smarter way to handle bad blocks used in PCs:
 - The controller maintains a list of bad blocks on the disk.
 - The list is initialized during the low-level format at the factory, and is updated over the life of the disk.
 - Low-level formatting sets aside spare sectors not visible to the operating system.
 - The controller can be told to replace each bad sector logically with one of the spare sectors.
 - This scheme is known as sector sparing or forwarding.

Bad Blocks (Cont.)

- A typical bad-sector transaction might be as follows:
 - The operating system tries to read logical block 87.
 - The controller calculates the ECC and finds that the sector is bad. It reports this finding to the operating system.
 - The next time that the system is rebooted, a special command is run to tell the controller to replace the bad sector with a spare.
 - After that, whenever the system requests logical block 87, the request is translated into the replacement sector's address by the controller.

Tertiary Storage Devices

- Low cost is the defining characteristic of tertiary storage.
- Generally, tertiary storage is built using *removable media*
- Common examples of removable media are floppy disks and CD-ROMs; other types are available.

Removable Disks

- Floppy disk — thin flexible disk coated with magnetic material, enclosed in a protective plastic case.
 - Most floppies hold about 1 MB; similar technology is used for removable disks that hold more than 1 GB.
 - Removable magnetic disks can be nearly as fast as hard disks, but they are at a greater risk of damage from exposure.

WORM Disks

- The data on read-write disks can be modified over and over.
- WORM (“Write Once, Read Many Times”) disks can be written only once.
- Thin aluminum film sandwiched between two glass or plastic platters.
- To write a bit, the drive uses a laser light to burn a small hole through the aluminum; information can be destroyed by not altered.
- Very durable and reliable.
- *Read Only* disks, such as CD-ROM and DVD, come from the factory with the data pre-recorded.

Speed

- Two aspects of speed in tertiary storage are bandwidth and latency.
- Bandwidth is measured in bytes per second.
 - Bandwidth – average data rate during a large transfer; number of bytes/transfer time.
- Access latency – amount of time needed to locate data.
 - Access time for a disk – move the arm to the selected cylinder and wait for the rotational latency; < 35 milliseconds.

Reliability and Cost

- A fixed disk drive is likely to be more reliable than a removable disk or tape drive.
- Main memory is much more expensive than disk storage
- The cost per megabyte of hard disk storage is competitive with magnetic tape if only one tape is used per drive.