Mass Storage

• Large storage capacity
• Non-volatile
• Much slower access
• Many technologies have been used
  • Magnetic
    • Drums
    • Tapes
    • Disks
  • Optical
    • CD R
    • CD rw
  • Solid state
Magnetic Devices

• Write by magnetizing an area using a recording head
  • Area must be moving with respect to the recording head
• Read by moving the recorded area against a read head
• Often use one head for reading and writing
• Movement must be stable
• Recording Density – bits/inch, or bits/sq inch
• Area
  • Drum – cylinder
  • Tape – a long ribbon
  • Disk – flat platter
Magnetic Drum
Magnetic Tape

• Was early secondary-storage medium
  • Evolved from open spools to cartridges
• Relatively permanent and holds large quantities of data
• Access time slow
• Random access ~1000 times slower than disk
• Mainly used for backup, storage of infrequently-used data, transfer medium between systems
• Kept in spool and wound or rewound past read-write head
• Once data under head, transfer rates comparable to disk
  • 140MB/sec and greater
• 200GB to 1.5TB typical storage
• Common technologies are LTO-{3,4,5} and T10000
The First Commercial Disk Drive

1956
IBM RAMDAC computer included the IBM Model 350 disk storage system

5M (7 bit) characters
50 x 24” platters
Access time = < 1 second
Mass Storage - Disk
Physical Structure of Recording Surfaces
Moving-head Disk Mechanism

• Multiple platters – One head per surface
• Track – Area that comes under a head from a surface
• Cylinder – In a multiplatter disk, the collection of tracks covered without moving the head assembly
• Sector – A track is organized into a collection of sectors. It is the unit of disk access.
Hard Disks

- **Magnetic disks** provide bulk of secondary storage of modern computers
  - Drives rotate at 60 to 250 times per second
  - **Transfer rate** is rate at which data flow between drive and computer
  - **Positioning time** (random-access time) is time to move disk arm to desired cylinder (seek time) and time for desired sector to rotate under the disk head (rotational latency)
  - **Head crash** results from disk head making contact with the disk surface -- That’s bad

- Disks can be removable
- Drive attached to computer via **I/O bus**
  - Busses vary, including EIDE, ATA, SATA, USB, Fibre Channel, SCSI, SAS, Firewire
  - **Host controller** in computer uses bus to talk to disk controller built into drive or storage array
Hard Disks

- Platters range from .85” to 14” (historically)
  - Commonly 3.5”, 2.5”, and 1.8”
- Range from 30GB to 16TB per drive
- Performance
  - Transfer Rate – theoretical – 6 Gb/sec
  - Effective Transfer Rate – real – 1Gb/sec
  - Seek time from 3ms to 12ms – 9ms common
  - Average seek time measured or calculated based on 1/3 of tracks
  - Latency based on spindle speed
    - $1 / (\text{RPM} / 60) = 60 / \text{RPM}$
    - Average latency = ½ latency

(From Wikipedia)
Hard Disk Performance

- **Access Latency** = **Average access time** = average seek time + average latency
  - For fastest disk 3ms + 2ms = 5ms
  - For slow disk 9ms + 5.56ms = 14.56ms

- **Average I/O time** = average access time + (amount to transfer / transfer rate) + controller overhead

- For example to transfer a 4KB block on a 7200 RPM disk with a 5ms average seek time, 1Gb/sec transfer rate with a .1ms controller overhead =
  - 5ms + 4.17ms + 0.1ms + transfer time = 9.27 ms
  - Transfer time = 4KB / 1Gb/s * 8Gb / GB * 1GB / 1024^2KB = 32 / (1024^2) = 0.031 ms
  - Average I/O time for 4KB block = 9.27ms + .031ms = 9.301ms
Solid-State Disks

• Nonvolatile memory used like a hard drive
  • Many technology variations
• Can be more reliable than HDDs
• More expensive per MB
• Maybe have shorter life span
• Less capacity
• But much faster
• Busses can be too slow -> connect directly to PCI for example
• No moving parts, so no seek time or rotational latency
Disk Structure

• Disk drives are addressed as large 1-dimensional arrays of **logical blocks**, where the logical block is the smallest unit of transfer
  • Low-level formatting creates **logical blocks** on physical media
• 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
  • Logical to physical address should be easy
    • Except for bad sectors
    • Non-constant # of sectors per track via constant angular velocity
Disk Attachment

- Host-attached storage accessed through I/O ports talking to I/O busses
- SCSI itself is a bus, up to 16 devices on one cable, **SCSI initiator** requests operation and **SCSI targets** perform tasks
  - Each target can have up to 8 **logical units** (disks attached to device controller)
- FC is high-speed serial architecture
  - Can be switched fabric with 24-bit address space – the basis of **storage area networks (SANs)** in which many hosts attach to many storage units
- I/O directed to bus ID, device ID, logical unit (LUN)
Storage Array

• Can just attach disks, or arrays of disks
• Storage Array has controller(s), provides features to attached host(s)
  • Ports to connect hosts to array
  • Memory, controlling software (sometimes NVRAM, etc)
  • A few to thousands of disks
  • RAID, hot spares, hot swap (discussed later)
  • Shared storage -> more efficiency
  • Features found in some file systems
    • Snapshots, clones, thin provisioning, replication, deduplication, etc
Storage Area Network

• Common in large storage environments
• Multiple hosts attached to multiple storage arrays - flexible
Storage Area Network (Cont.)

- SAN is one or more storage arrays
  - Connected to one or more Fibre Channel switches
- Hosts also attach to the switches
- Storage made available via LUN Masking from specific arrays to specific servers
- Easy to add or remove storage, add new host and allocate it storage
  - Over low-latency Fibre Channel fabric
- Why have separate storage networks and communications networks?
  - Consider iSCSI, FCOE
Network-Attached Storage

- Network-attached storage (NAS) is storage made available over a network rather than over a local connection (such as a bus)
  - Remotely attaching to file systems
- NFS and CIFS are common protocols
- Implemented via remote procedure calls (RPCs) between host and storage over typically TCP or UDP on IP network
- **iSCSI** protocol uses IP network to carry the SCSI protocol
  - Remotely attaching to devices (blocks)