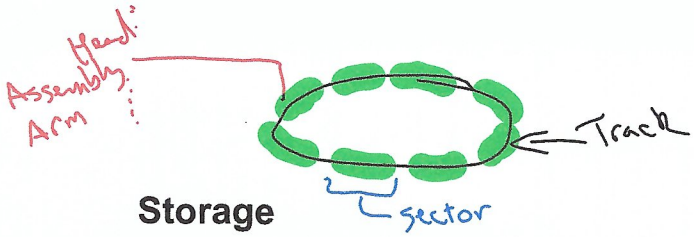
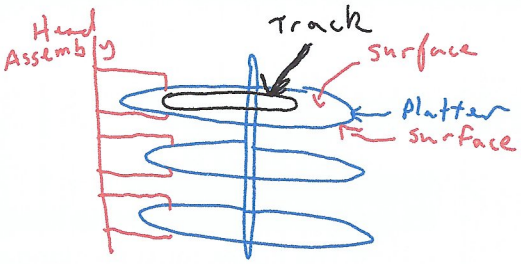
Memory → Storage → Fault Tolerance





Storage

Storage is slower with a higher capacity than main memory.
Performance of storage is based on two characteristics:

throughput → which is improving — throughput is how many bytes per second we can move in/out of storage

latency → which is improving very slowly — After we request a page of data, latency is how long until we get it back.

Storage reliability is very important. If processor fails, we can replace it and reboot. If storage fails, we lose everything.
There are many diverse types of storage.

Magnetic Disks

Disk Capacity = (# of surfaces)(# of tracks / 1 surface)(# of sectors / 1 track)(# of bytes / 1 sector)

(# of surfaces) = # platters x 2

(# of tracks / 1 surface) = # of cylinders on a surface

Magnetic disks = Hard drives and floppy disks
Data bits are on both surfaces of the disks.
The "cylinder" or "track" of a surface is determined by the radial location of the head assembly.

Access Time for Magnetic Disks

Seek time - time to move head assembly to the correct cylinder
Rotational latency - time to move to disk to the correct sector
Data Read - time required to read the entire sector
Controller time - time to check data (check sum, etc)
I/O Bus time - time to load data on the bus and send to processor

Only one disk access can be performed at a time, so there is an additional
Queuing delay - time to wait for availability of the disk (aka waiting for previous disk accesses to complete before yours is serviced).
We can only read one track at a time. We cannot move to another track until we're done reading from the first track.

Trends for Magnetic Disks

Capacity is improving — 2x every 1-2 years
Seek Time has slowly improved — Can improve seek time by making a faster, more accurate head assembly motor (which is hard), or by making the disks smaller and more compact (which is what we've been doing).
Rotation Speed has improved
Controller and Bus Speed have improved
Disk access time is mostly dominated by the Seek Time and (to a lesser extent) the Rotation Latency.

Optical Disks

Optical disks (such as CDs and DVDs) are similar to magnetic disks. The main difference is that they use lasers instead of head assemblies.
Since they are portable, they need to be standardized, which slows improvement.

Magnetic Tape

Magnetic tape is used for secondary (backup) storage.
They have a large capacity, but must be read sequentially.

Using RAM for Storage

RAM is benefitting from Moore's Law. As opposed to Hard Drives, which don't benefit from Moore's Law.
SSD = Solid State Disk — Not a "disk" at all. All electronic storage.

Two options for SSD:
- SSD can be a DRAM and Battery, it is fast, but expensive.
- Flash Memory uses transistors and is low power. Flash is significantly faster than Disk, but slower than DRAM. Can keep data alive without power.

Hard Drives are also very slow compared to RAM (especially DRAM).
RAM is improving capacity due to Moore's Law (transistors getting smaller), but they're significantly faster than Hard Drives. HDD access times are about 10 ms whereas RAM access times are well under a microsecond (3 orders of magnitude difference).
However, we use Disk because it is 100x cheaper per GB, even though DRAM has about 100,000x better latency.

Hybrid Magnetic Flash

A combination of magnetic disk and flash drive. The flash drive is used as a cache for the disk.

Magnetic Disk Pros:Low cost per GB of storage and we know how to build them with Huge Capacity in a relatively small package.
Magnetic Disk Cons: It is very power hungry (due to spinning of disk, even when not reading data), it is slow due to needing to wait for mechanical movements, & it is sensitive to impact.

Flash Pros: Fast, Power Efficient, & no moving parts.

Connecting IO Devices

IO devices are connected to the system with a bus.

The buses must be standardized to connect to a number of IO devices, so improvement to the bus are slow.

Storage (& other I/O devices) connect via Bus. Buses need to be standardized, but that leads to slow improvement adoption.
Computers have a Hierarchy of Buses;
- Mezzanine Bus: (ex: PCI Express). This is the fastest bus that connects directly to the processor. Other fast devices (the graphics) connect to this bus.
- SATA/SCSI Bus: This is a slower bus that we connect our Storage Devices to (which all understand the "standardized" storage interface). There is a "SATA Controller" that connects this bus to the Mezzanine (PCI Express) bus. This separate, slower bus allows us to separate the slower drives from the fast devices, allowing updates to the fast devices and communication standards to improve more rapidly while the slower Storage Devices can keep their own old standardizations for longer. If there was an update to PCI Express (for example), we'd just need to update the SATA controller- the Storage Devices themselves can stay the same.
- USB Bus: Another separate bus that connects to the Mezzanine Bus. Same reason as the SATA/SCSI Bus.