Solid State Drives

Leaving disks in the dust
Memory Hierarchy

- Computer memory consists of three major categories.
  - Primary
    - Internal
      - Registers-
        - Accessed in one cycle
        - Size in hundreds of bytes range.
      - Level 1 cache-
        - Accessed in a few cycles
        - Size in tens of kilobyte range
      - Level 2 cache-
        - Access time 2x - 5x longer than L1 cache
        - Size 512 kiB or more
      - Level 3 cache-
        - Higher latency than L2
        - Size 2048 kiB or more
Memory Hierarchy (Continued)

- Main Memory
  - Random Access Memory (RAM)
    - Speed varies with type of RAM
      - Types consist of SRAM, DRAM, PRAM
    - Computers can have from 8-16 GB or more
  - Secondary
    - Disk
      - Consumer solid state drives (SSD) operate at 100 - 600 MB/s
      - Size of consumer SSDs ranges from 64GB-2TB
      - Consumer hard disk drives (HDD) operate at 140MB/s
      - Size of consumer HDDs ranges from 1 TB-4 TB
Memory Hierarchy (Continued)

- Tertiary
  - Tape libraries-
    - Size ranges from 20 TB - 411 PB
    - Access time varies from seconds to minutes.
  - Optical jukeboxes-
    - Size ranges from TB to PB
    - Access time 4-9 seconds depending on disk swaps.
Memory Hierarchy (Continued)

Computer Memory Hierarchy

- Power on:
  - Immediate term: small size, small capacity
  - Very short term: medium size, medium capacity
  - Short term: small size, large capacity
  - Mid term: large size, very large capacity
  - Long term: large size, very large capacity

- Power off:
  - Flash/USB memory: random access memory
  - Slow, cheap:
    - Hard drives: fast, affordable
  - Very slow, affordable:
    - Tape backup
Characteristics of Memory (Volatility)

- Volatility
  - **Non-volatile**: Retains storage data even if power is not constantly supplied to unit.
  - **Volatile**: Requires constant power to maintain memory.
  - **Dynamic Random Access Memory**: Form of volatile memory which must be periodically refreshed or it will be lost.
  - **Static Random Access Memory**: Similar to DRAM except it does not need to be refreshed as long as it maintains a power source.
Characteristics of Memory (Mutability)

• Mutability
  o **Read/Write Storage** - Capable of overwriting information at any time. Modern computers use read/write storage in secondary memory.
  o **Read Only Storage** - Also known as immutable storage. Read only storage allows only one write typically at the time of manufacturing. It is featured in tertiary memory.
  o **Slow Write, Fast Read Storage** - Allows multiple overwrites, but has a much slower read than write process. Used in some flash memory.
Characteristics of Memory (Accessibility)

- Accessibility
  - Random Access - Any piece of memory may be accessed at any point. The ideal access time for this type of memory is constant however that is not always the case. This type of accessibility is often used in many primary and secondary types of storage.
  - Sequential Access - Memory may only be accessed in a serial order. This type of accessibility is generally only used in offline types of storage.
Characteristics of Memory (Addressability)

- **Addressability**
  - **Location Addressable** - Each unit of accessible information is stored with an associative memory address. Not user friendly due to non-human readable format.
  - **File Addressable** - Information is divided into file of variable size. These files are then presented to users in a human readable format with directories and filenames.
  - **Content Addressable** - Each unit of accessible information is selected on basis of the contents stored there.
Characteristics of Memory (Capacity)

- Capacity
  - **Raw Capacity** - Total amount of information the storage device may hold. Defined in units of bits and bytes.
  - **Memory Storage Density** - How compact information is stored within the device. Comes in units of memory per space, ex MegaBytes per square inch.
Characteristics of Memory *(Performance)*

- **Performance**
  - **Latency**: The time necessary to access a specific location in storage. The units of measurement run from nanoseconds, milliseconds, and seconds depending on the layer of memory.
  - **Throughput**: The rate at which information may be written to or read from storage. Read rates and write rates may vary. Often measured in MB/S.
  - **Reliability**: Probability a bit value may change under environmental conditions.
HDD Primary Components

- Components
  - **Platter** - Magnetic disk shaped object, which holds data. Consumer platters are spun at speeds of 5,400 to 7,200 rotations per minute.
  - **Read/Write Head** - Used to detect and modify magnetic fields on the spinning platter.
How an HDD operates

• The platter (a magnetic disk) is spun at very high speeds by a series of motors. An arm is used to position the head over the location where data needs to be written or read. The head then detects a magnetic value during a read, or modifies the magnetic value during a write.

• Due to using a magnetic disk, HDDs are very vulnerable to magnets.
SDD Primary Components

• Components
  o **Controller**- Incorporates electronics used to bridge NAND memory components to host computer. It is an embedded processor which executes firmware level code. Controllers contribute to a large percent of the performance of a solid state drive.
  o **Memory**- Uses integrated circuits to store and persist data. Currently most manufactures use NAND flash based memory, due to lower cost when compared to dynamic random access memory and because NAND flash memory is nonvolatile. DRAM solutions however may be faster than flash based SSDs
How a SSD Operates

• Unlike a HDD, solid state drives do not contain and solid mechanical moving parts.
• This makes SDD more reliable in regards to mechanical failure.
• The host computer sends data to the SSD controller which then decides how to write the data to the NAND memory. This is done similar to RAID, as the controller writes to multiple NAND chips at once.
**History of SSDs**

- Concept of solid state drive originated during the 1950's.
- Developed same time as magnetic core memory and card-capacitor read only store.
- Unfortunately, technology was abandoned temporarily for much cheaper drum storage unit.
- SSDs became more popular again in the 1970's due to being implemented with semiconductors. IBM, CRAY, and Amdahl used SSDs to create some of their earlier...
• Solid state drives were however forgotten about again due to high costs.

• "In 1978, Texas Memory Systems introduced a 16 kilobyte RAM solid drive to be used by old companies for seismic data acquisition".

• The next year, "StorageTek developed the first RAM solid state drive."

• In 1986, "Santa Clara Systems introduced BatRam, a 4 megabyte mass storage system expandable to 20MB using 4MB memory modules." It was particularly innovative because it included a rechargeable battery, which protected the memory chip from losing its contents when power was lost.
• In 2009 RAM solutions were still being used because they were faster than the fastest solid state drives, however RAM solutions used more CPU resources.

• "In 1995, M-Systems introduced flash-based solid-state drives" These drive were innovative because they did not require batteries to maintain memory.

• However at this time flash based SSDs were still slower than DRAM based solutions.
Since then, SSD solutions have been used in the military and many other industries. This is due to being nonvolatile and having an exceptional mean time between crashes.

SSDs can also withstand extreme vibrations, shocks, and temperatures.
Within the last twenty years large leaps have been made in the flash-based SSD industry.

In 2007 Fusion-io announced a Peripheral Component InterConnect Express (PCIe) based solid state disk capable of performing 100,000 input/output operations per second or IOPS.

Traditionally SSDs consisted of DRAM, but in 2010 the industry switched to NAND-based flash memory, which retains data without power.
Solid State Drives Today

• Currently Virident Systems has produced a flash based SSD with a form factor of $\frac{1}{2}$ length x $\frac{1}{2}$ height capable of delivering a read bandwidth up to 2.7 gigabytes per second and 1.5 million IOPS in a single low-profile card.

• Since 1994, the size of SSD’s has increased from eighteen gigabytes to an incredible 2.2 terabytes, which is an amazing 12222.2 percent increase.
Solid State Drives Tomorrow

- Researchers from Microsoft predict the size of SSD will be in neighborhood of sixteen terabytes in next decade.
- Researchers have also linked higher error rates and latency with the increased size of SSDs.
- Studies have illustrated a sixteen terabyte SSD will be too unstable for a reliable performance.
Solid State Drives Tomorrow (Continued)

- Also believe speed advantages observed by using SSDs will disappear due to increased latencies.
- Some analysts claim 3D memory may be the successor to flash memory. 3D memory is currently under development and not commercially available.
Comparing Solid State Drives to Hard Disk Drives

- Often times it can be difficult to compare the features of a SSD and a HDD. This is because SSD benchmarks measure values which HDD's have traditionally performed poorly at.
- Two such values are seek time and rotational latency. Because SSD do not have to spin in order to acquire data, they are vastly superior in these fields.
Comparing SSDs to HDDs (Continued)

- SSD startup almost instantaneously due to lack of moving mechanical parts. Could take a few ms to come out of power saving mode.
- Hard disk drives take much longer to start up. Contributed to disk spin up.
- Spin up refers to the process of the hard disk accelerating its platters from stopped state to an operational speed.
- Computers with multiple hard disks may have to stagger spin up cycles in order to limit the amount of power draw.
Comparing SSDs to HDDs (Continued)

- Random access time for SSD < is 100 microseconds
- Data can retrieved directly from various locations in flash memory.
- HDD access time ranges from 2.9 to 12ms.
- Large access time attributed moving head rotation speed.
- SSD read latency time is lower. Can read from any location within drive.
- HDD latency is considerably higher.
Comparing SSDs to HDDs (Continued)

• Data transfer rate for reads/writes on SSD very consistent.
• Performance reduced when large number of individual blocks are accessed.
• Consumer SSD transfer rate = 100 MB/s - 600 MB/s
• Enterprise SSD transfer rate = multi GB/s
• HDD takes time to get head positioned.
• Once positioned enterprise level can transfer 140 MB/s In practice however could be less.
Comparing SSDs to HDDs (Continued)

- Blocks greater than FS not necessary for SSD to perform optimally.
- Defragmenting SSD is actually harmful due to limited number of reads and writes.
- HDD performs better when its files are not fragmented (Header must travel to multiple locations).
- HDD performs better after defragmentation
Comparing SSDs to HDDs (Continued)

- SSDs generally do not require additional cooling.
- Tolerate higher temperatures.
- Seagate estimates HDD reliability compromised past 55 °C.
- Seagate estimates HDD life shortened past 35 °C.
- SDD cost $0.59 per GB while HDD cost $0.05 per GB.
Comparing SSDs to HDDs (Continued)

- SSD less noisy than their hard disk cousins.
- SSD less susceptible to shock and vibrations.
- SSD physically more reliable, but limited number of writes/reads. DRAM solutions however do not have this issue.
- SSD data not damaged by magnets.
Flash Memory

• NOR Flash
  o Each row is wired together as well as each column.
  o Allows individual bits to be read, applying voltage to a word line allows the bit lines to show a voltage based on the charge that is stored.
  o Chips are complex as the word and bit line connections take a large amount of space.
• NAND Flash
  o Word lines same as NOR flash, bit lines are connected in series. This allows the chip to be 60% smaller than an equivalent NOR chip.
  o All bit lines deliver a signal simultaneously, in parallel.
  o Reading and writing is done blocks at a time, but this is how hard drives function.
Flash Memory

(Continued)

• Single level cell
  o Stores a single bit of information.
  o More reliable, lower error rate.
  o Last for more write cycles.
  o Usually found at the enterprise level.

• Multi level cell
  o Can store multiple bits in a single cell.
  o More complex, less reliable.
  o Fewer write cycles.
  o Lower cost, dominates the consumer market.
Limitations

- Erasing data can only be done a block at a time.
- Finite number of program-erase cycles, usually at least 100,000.
- Writing to one cell multiple times can force surrounding cells to change if they are not rewritten periodically.
- Individual flash cells are slow.
Memory Controllers

- Perform various tricks to increase performance:
  - error correction
  - wear leveling
  - bad block mapping
  - garbage collection
  - read and write caching
  - read disturb management.
  - read from multiple NAND chips at once, similar to RAID.
• **Error Correction** - performed by writing and reading data in stripes across different NAND chips, this effectively makes SSDs a RAID array of NAND. Some controllers also provide parity protection.

• **Wear Leveling** - keep track of the writes and dynamically remapping block in order to spread write operations between sectors.

• **Bad Block Management** - perform write verification and remap to spare sectors in case of write failure.
Memory Controllers (Continued)

- Garbage Collection -
  - As pages are rewritten, the controller keeps track of which pages contain good data and which have stale data. Stale pages are not doing anything but cannot be individually erased.
  - During downtime the controller will take an entire block of pages, copy the good data to a different block and then erase the entire first block.
  - Constant garbage collection can have significant effects on the lifespan of an SSD.
Memory Controllers (Continued)

- Read and write caching
  - Some SSDs incorporate SDRAM as part of their storage, this cache allows the SSD to quickly receive data that it needs to write, the data sits in the SDRAM until the controller has time to actually commit it to NAND.
  - This greatly decreases latency and increases throughput, but the SDRAM will erase itself if it loses power. Usually a large capacitor is used to power the drive should power be lost.
  - SDRAM is usually between 128 and 512 MB.
Read Disturb Management - Track number of reads across the storage device and rewrite the surrounding data periodically.
References

- http://en.wikipedia.org/wiki/Memory_hierarchy