Data storage has various characteristics

Addressability
Word-level addressing—Very efficient computationally. Hardware intensive.

File addressable—There is an "allocation table" that makes use of file names and possibly other hierarchy such as directories or folders. Optionally (usually), files may be accessed in parts, usually called "sectors." The allocation table must be used to find the start of the file. Then the parts of the file are accessed by the sector offset from the start of the file.

Content-addressable memory (CAM)—a chunk of reserved memory that operates as a buffer in sort of a reverse telephone directory mechanism. Sometimes called a translation lookaside buffer (TLB) or an associative memory. The CPU supplies a word of information—typically a virtual address from which data is desired. The CAM finds the matching word and returns the physical address at which the desired data may be found. Essentially the CAM does a search via hardware, not software. This type of memory is essential to a fast cache system.

Example: CAM used to find a rough estimate of log(x). The numbers 1, 10, 100, 1000, etc are stored in CAM at addresses 0, 1, 2, 3, etc. Suppose the CPU supplies the number 150 to the CAM. Address 2 contains 100 and address three contains 1000. The CAM searches and nearly instantly finds that address 2 is the last address containing content less than 150. The content addressable memory returns data of 2. Typically this is then used as an initial guess for an iterative solution of the desired precision.

Note: This can also be done with a software algorithm using conventional memory. (slow!)

1

Data storage has various characteristics

Capacity
Total amount stored. bits or Bytes.
Note: k, M, T, G. etc are base 10 prefixes.
Ki, Mi, Ti, Gi, etc. are base 2 prefixes.
e.g. 1 kb = 1000 bits,
1 kB = 1000 bytes,
1 kiB = 1024 bytes.

Value		prefix
10^3 = 1 000	k	kilo
10^6 = 1 000 000	M	mega
10^9 = 1 000 000 000	G	giga
10^{12} = 1 000 000 000 000	T	tera
10^{15} = 1 000 000 000 000 000	P	peta
10^{18} = 1 000 000 000 000 000 000	E	exa

Value		prefix
2^{10} = 1 024	ki	kibi
2^{20} = 1 048 576	Mi	mebi
2^{30} = 1 073 741 824	Gi	gibi
2^{40} = 1 099 511 627 776	Ti	tebi
2^{50} = 1 125 899 906 842 624	Pi	pebi
2^{60} = 1 152 921 504 606 846 976	Ei	exbi

Density
Bits per unit volume, area, mass, or weight.
Certain technologies store more bits in less physical space.

Organization
Number of number of words x number of bits/word
e.g. 4 GB x 32 is the organization of a 128 GB memory

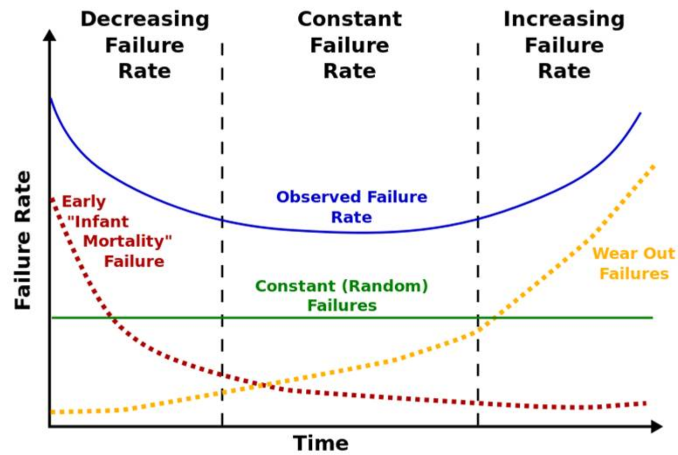
Speed
Latency—time from request for data to delivery of data
nanoseconds for primary storage
not so important for most other memories

Throughput—maximum sustained rate of data delivery
Usually this needs to be high for all media

2

Data storage has various characteristics

Reliability
The probability of a spontaneous bit change or random breakdown.
Often expressed as "mean time between failures" or MTBF.
The reciprocal is the "failure rate," λ which is sometimes used instead.
This is the green line in the illustration below.



https://commons.wikimedia.org/wiki/File:Bathtub_curve.svg

3

Data storage has various characteristics

Power consumption per bit stored
The power required to store a bit varies with the technology. Non-volatile storage appears to cost no power for long-term storage, but this is not true. The storage needs to be maintained—backed up at the least.
In addition to maintaining the data, reading and writing require power and must be done to utilize the data.
Various estimates are that about 2% of all electricity generated is used for data storage operations.
In the U.S. alone data storage consumes about 70 TWh/yr http://eta-publications.lbl.gov/sites/default/files/lbnl-1005775_v2.pdf
Power efficiency of data storage is now an area of active research
Power efficiency matters for battery life and for pollution reduction—a very real environmental issue.

Three business models are emerging for archival storage
1. Rental—an example is Amazon's Simple Storage Service (S3), $0.023/GB paid monthly. (has ~doubled in last 3 years)
2. Monetization of information—an example is a free gmail account
3. Endowment—a fund of money is established up front or via periodic payments for a limited time to support storage of data "forever." As of now it costs about $5/GB to store large quantitates of data "forever." (Large quantities means at least several terabytes.) https://www.cni.org/wp-content/uploads/2011/08/cni_nsf_goldstein.ppt

Example of "personal" endowed storage: $200 one-time-payment stores 10 GB "forever." ($20/GB)
https://www.forever.com/forever_storage

~~4. A thumb drive in a safety deposit box. This is not considered archival. Flash memory fades with time. Someone needs to keep up payment on the safety deposit box. Flash memory chip might just spontaneously fail. . .~~

4

Data storage media

Semiconductor

Static RAM. Usually organized word-wide, volatile, applications to main memory and cache memory
Dynamic RAM. Usually sold in modules but the internal chips are bit- or nibble-wide. Requires refresh.
ROM. Read Only Memory. Available in several sub-varieties
--mask programmed at the factory. Unchangable after that. Changes require physical replacement.
--EPROM Electrically Programmable Read Only Memory. One-time programmable by customer
Some EPROMs can be erased entirely by exposure to ultraviolet light.
--EEPROM Electrically Erasable Read Only Memory. Can be erased, re-written one word at a time.
--Flash memory. Same as EEPROM except can only be erased in large blocks of data
(Most "solid state hard drives" are flash memory combined with a RAM cache.)

Issues: All semiconductor technologies can be disturbed by light, heat, and x-rays.
RAM is subject to errors caused by nuclear radiation, which is naturally present in low amounts.

All EPROM, EEPROM and Flash technologies use "floating gate technology" which is subject to a slow decay of the stored electronic charge. Usually the expected reliable interval of storage is in the range of 10 to 35 years. (A lot of electronic devices are going to be useless when they are antiques.)

All EPROM, EEPROM, and Flash technologies use "floating gate technology" which suffers "wear." There is a maximum number of allowable write cycles. For EPROM and EEPROM this can be as low as 10 or 100 writes. Most Flash memories can withstand between 10^4 and 10^6 writes /word.

5

Data storage media

Semiconductor

Ferroelectric RAM (F-RAM, FeRAM, FRAM, etc.)
Acts the same as Flash memory except faster writes, much more endurance, > 10^{10} write cycles.
Disadvantage: high cost, low density

Magnetic

Floppy disk 1.44 MB/disk is cheap. About 1 GB/disk is possible, but never was popular in the marketplace.
Hard disk—several TB possible in one unit. Very cost effective for secondary storage, but "solid state. . ."
Tape reel-to-reel—obsolete
cartridge—in combination with "juke-box" changer allows truly massive tertiary storage
200 GB/square inch of thin tape is current state-of-the-art, and improvements are still coming.
https://spectrum.ieee.org/computing/hardware/why-the-future-of-data-storage-is-still-magnetic-tape
Drum obsolete—a hard disk on a cylinder instead
Core memory—small iron donuts that can be magnetized to store information—main memory on moon shots
All but tape cartridge systems are obsolete.

Magneto-optical

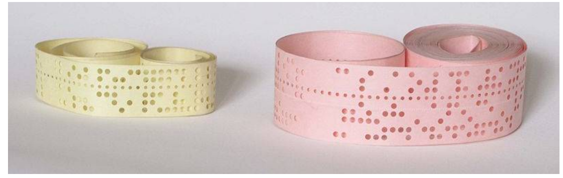
Write using magnetism and optics, read optically. Slow-write, fast read offline storage. (Too little too late.)

Optical

CD, CD-ROM, DVD, BD-ROM Read only offline storage
CD-R, DVD-R, BD-R, etc write once offline storage
CD-RW, DVD-RW, DVD+RW, DVD-RAM, BD-RE, etc slow write, fast read offline storage
All optical technologies are entering obsolescence

6

Obsolete Data storage media



Paper

Paper tape
Printed text with optical character recognition.

Vacuum Tube

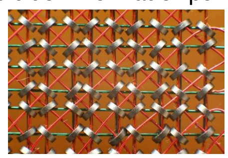
Williams tube, Selectron Tube. Both were fast primary storage for their era.

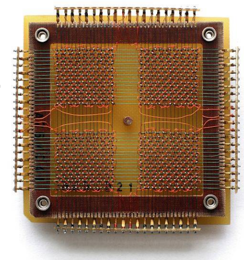
Electro – acoustic

Sound waves in a media with slow velocity of transmission, e.g. Hg.
Repeat the data or change it with every echo.

Core Memory

Ferrite donuts, each the size of a pin-head, stores one bit of information per donut.
Bit is stored as the polarization of a magnetic field.
https://en.wikipedia.org/wiki/Magnetic-core_memory





Futuristic data storage technology
DNA > 100 PB/gram has been achieved in the lab

https://commons.wikimedia.org/wiki/File:KL_Kernspeicher_Makro_1.jpg
https://commons.wikimedia.org/wiki/File:KL_CoreMemory.jpg
https://commons.wikimedia.org/wiki/File:PaperTapes-5and8Hole.jpg

7