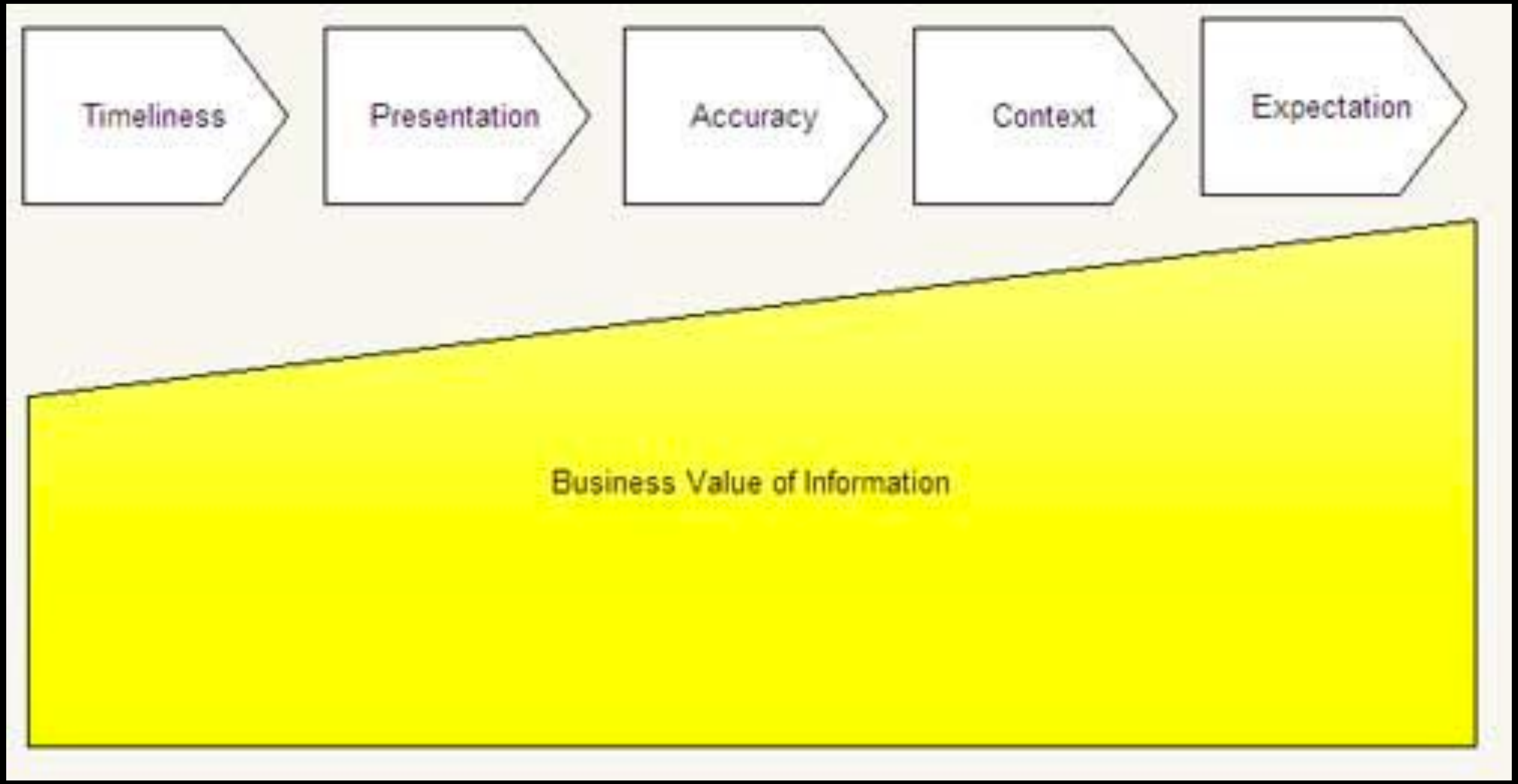
# Äriinfosüsteemid
# Business Information Systems

**Heino Talvik**
heino.talvik@gmail.com

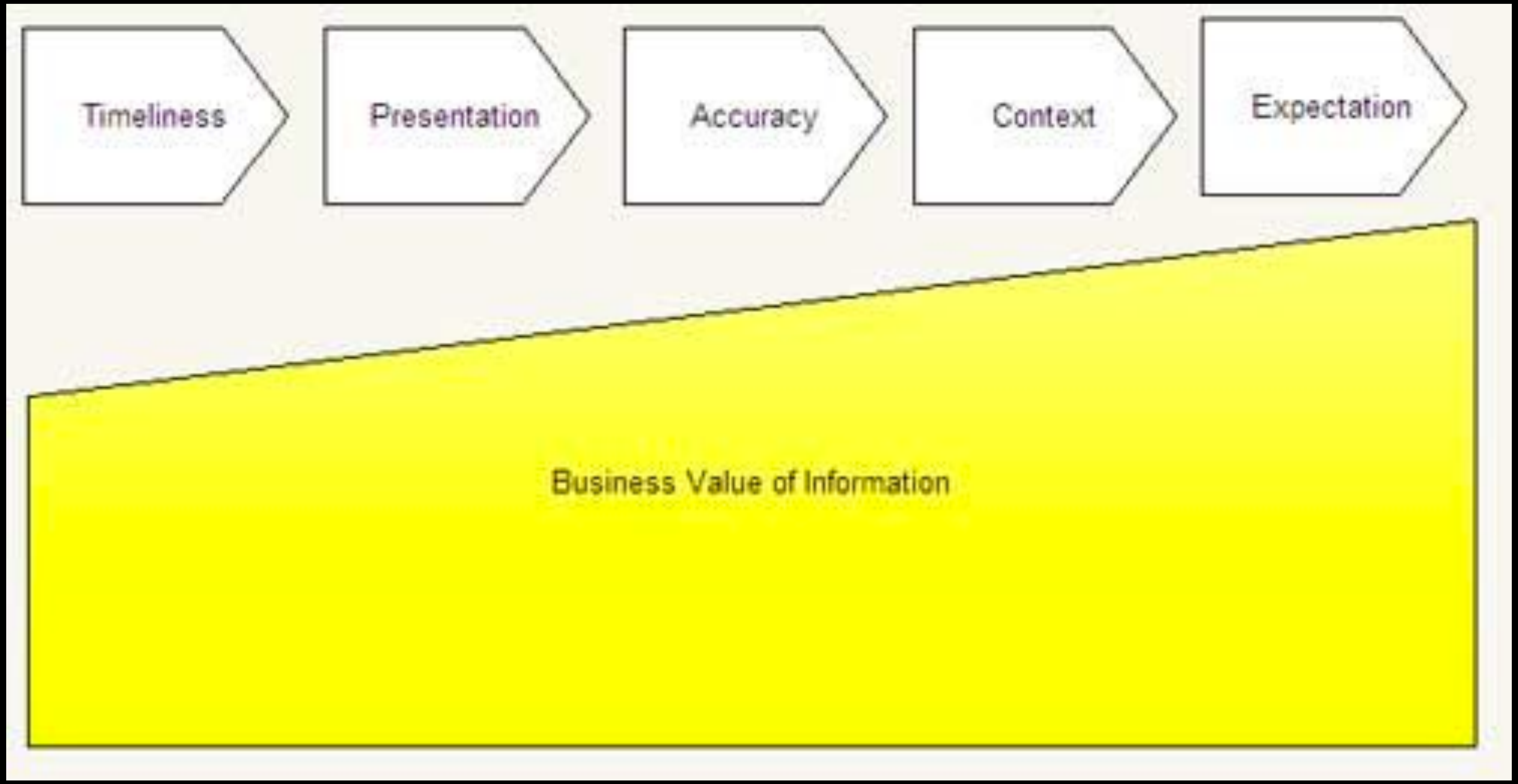
# Purpose of Business Information System

… getting the right information to the right people, at the right time, in the right amount, and in the right format.



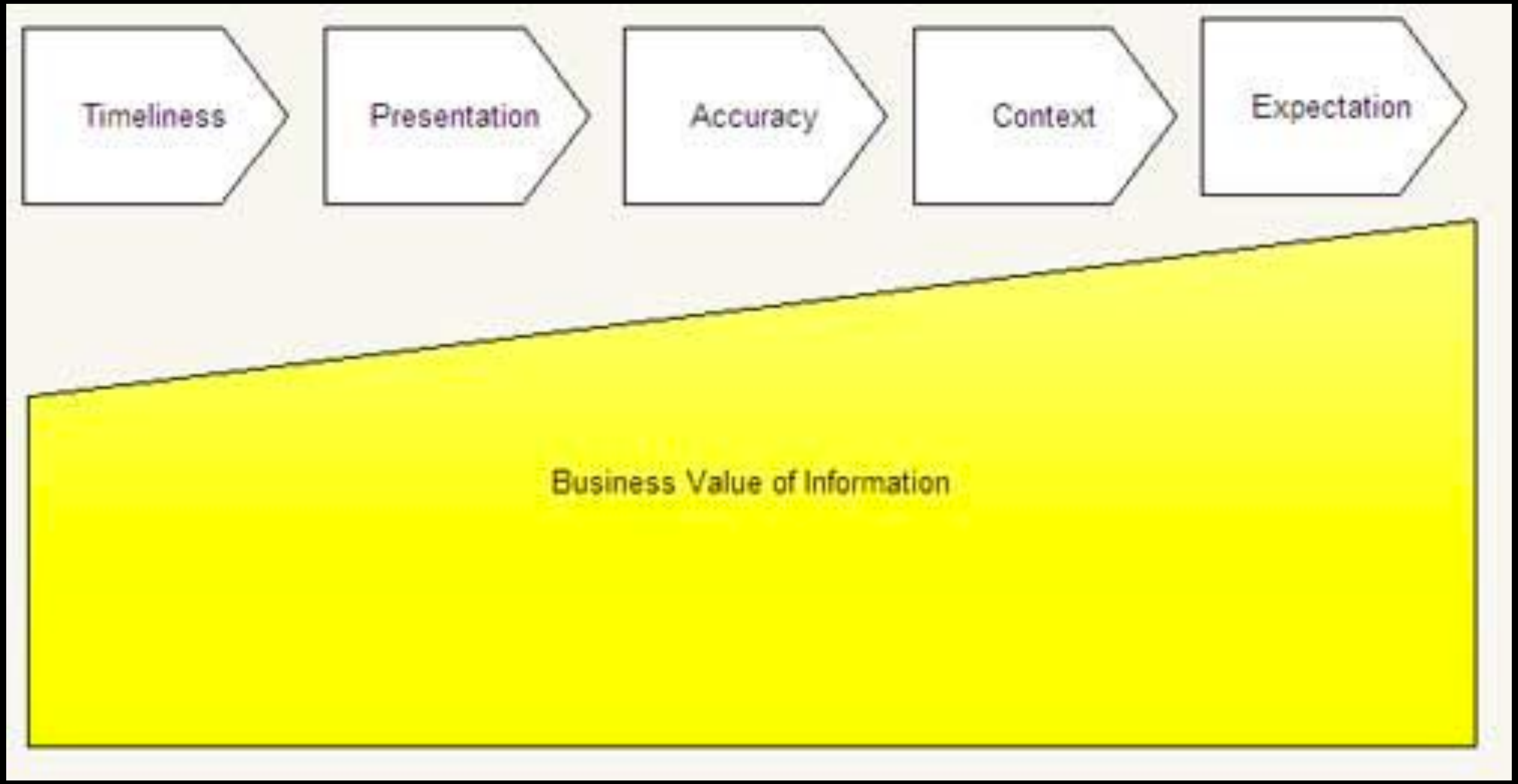
# Purpose of ……………………………………………… System

… getting the right …………………… to the right people, at the right time, in the right amount, and in the right format.



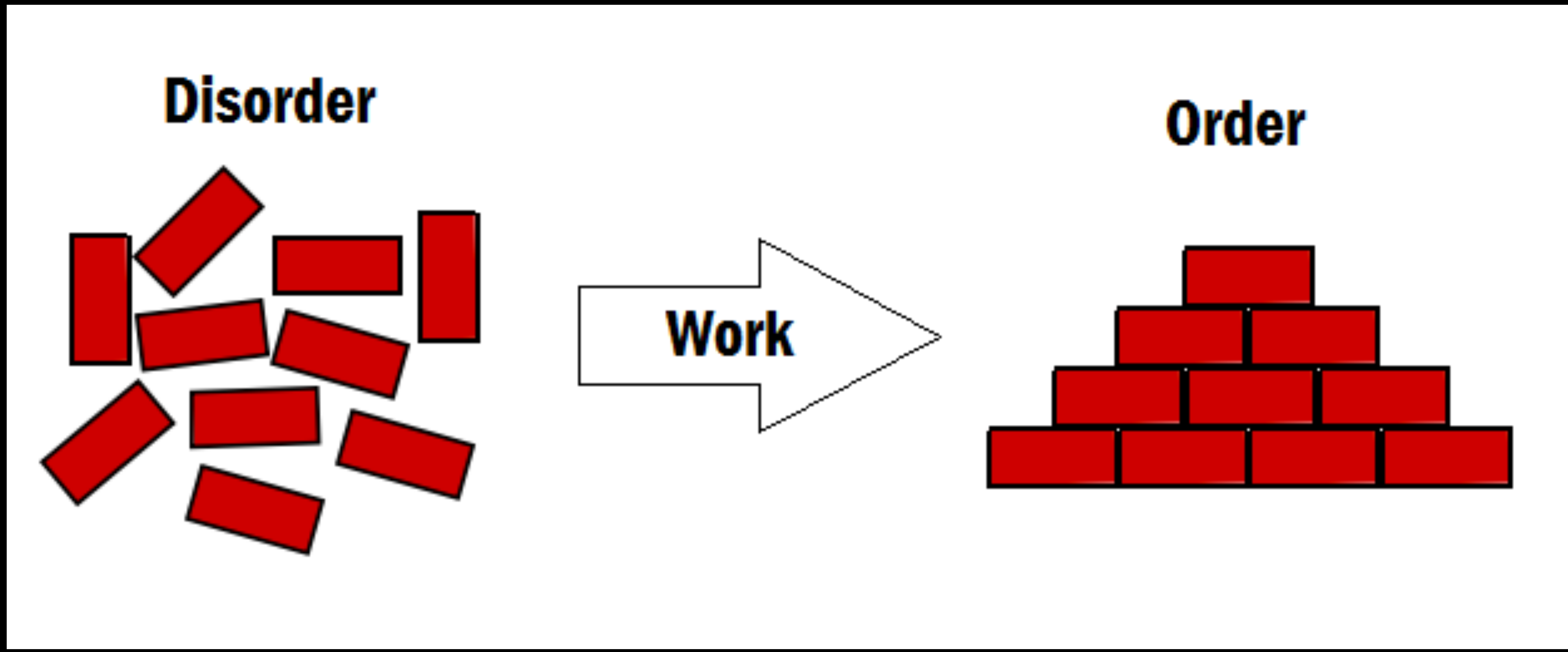
# Purpose of ………………………………………………………………… System

… getting the right ……………………………. ı to the right people, at the right time, in the right amount, and in the right format.



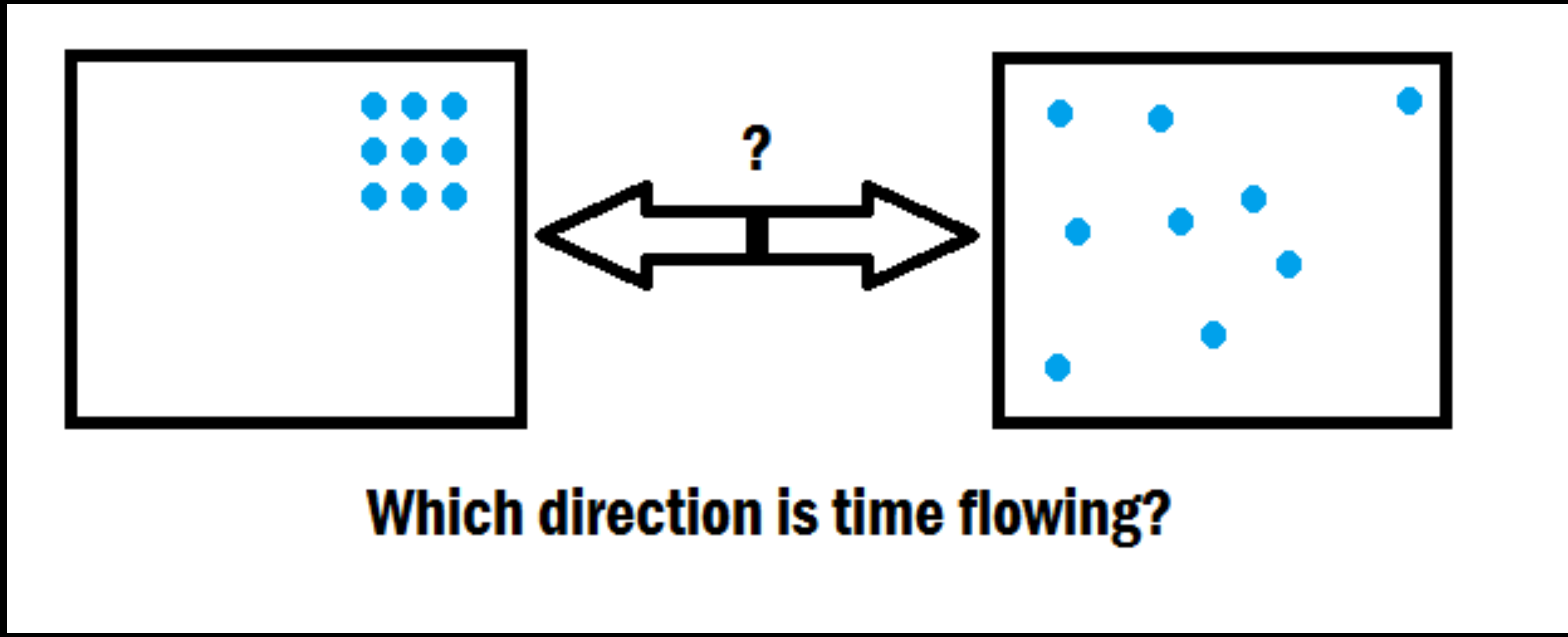
# Entropy

- … represents the degree of disorder or randomness in a system, and the tendency for all matter and energy in the universe to become evenly distributed.
- Second Law of Thermodynamics states: "the entropy of a closed system will never decrease over time, but will instead tend towards a maximum value”.



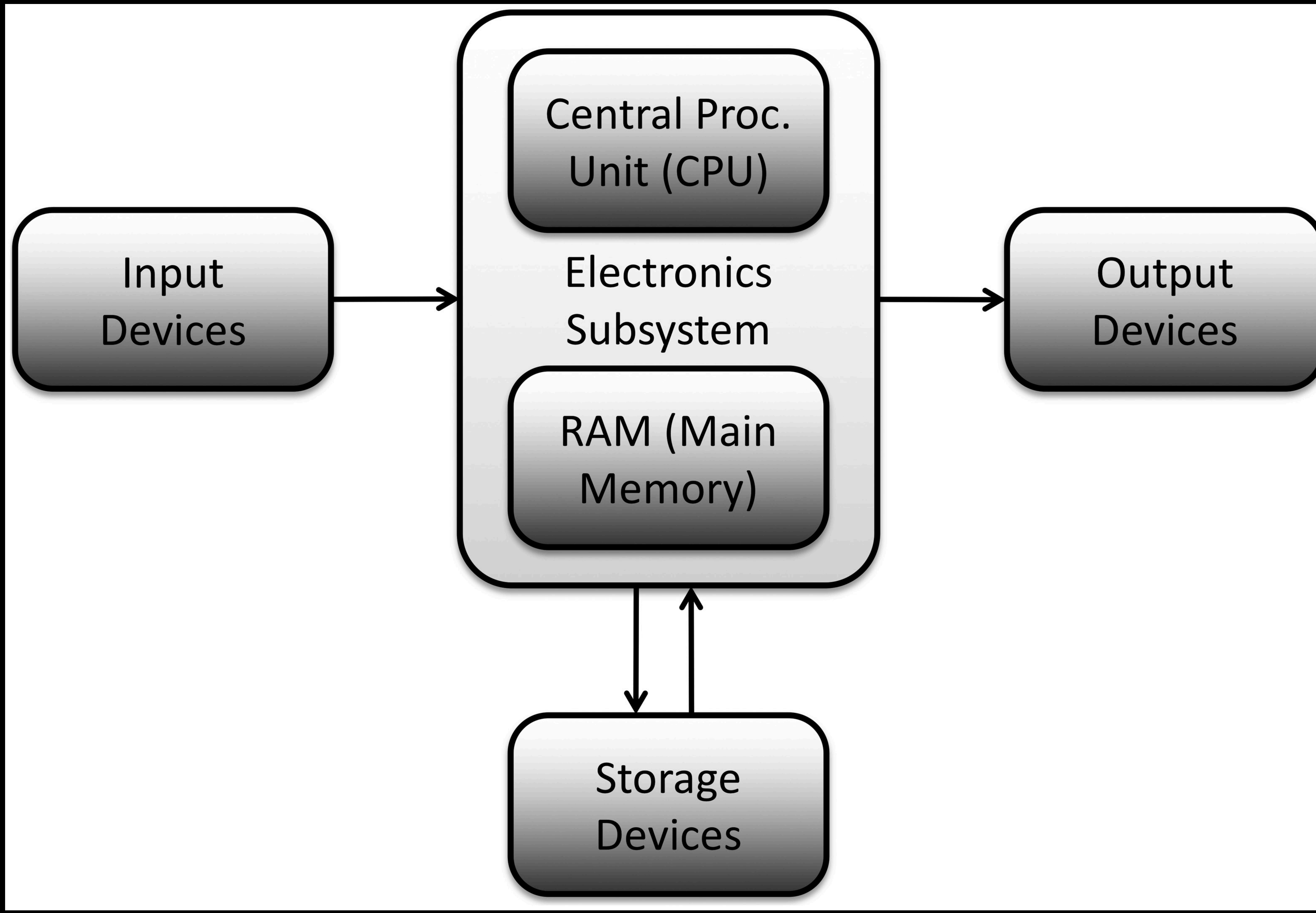
# Entropy in Information Systems

- … is used to measure the randomness or uncertainty of information in a digital signal, file, or data storage system.
- In information theory, entropy is a fundamental concept used to quantify the amount of information contained in a message or signal.
- Higher entropy signals have more unpredictability and randomness, while lower entropy signals have more order and structure.



# Hardware – computer

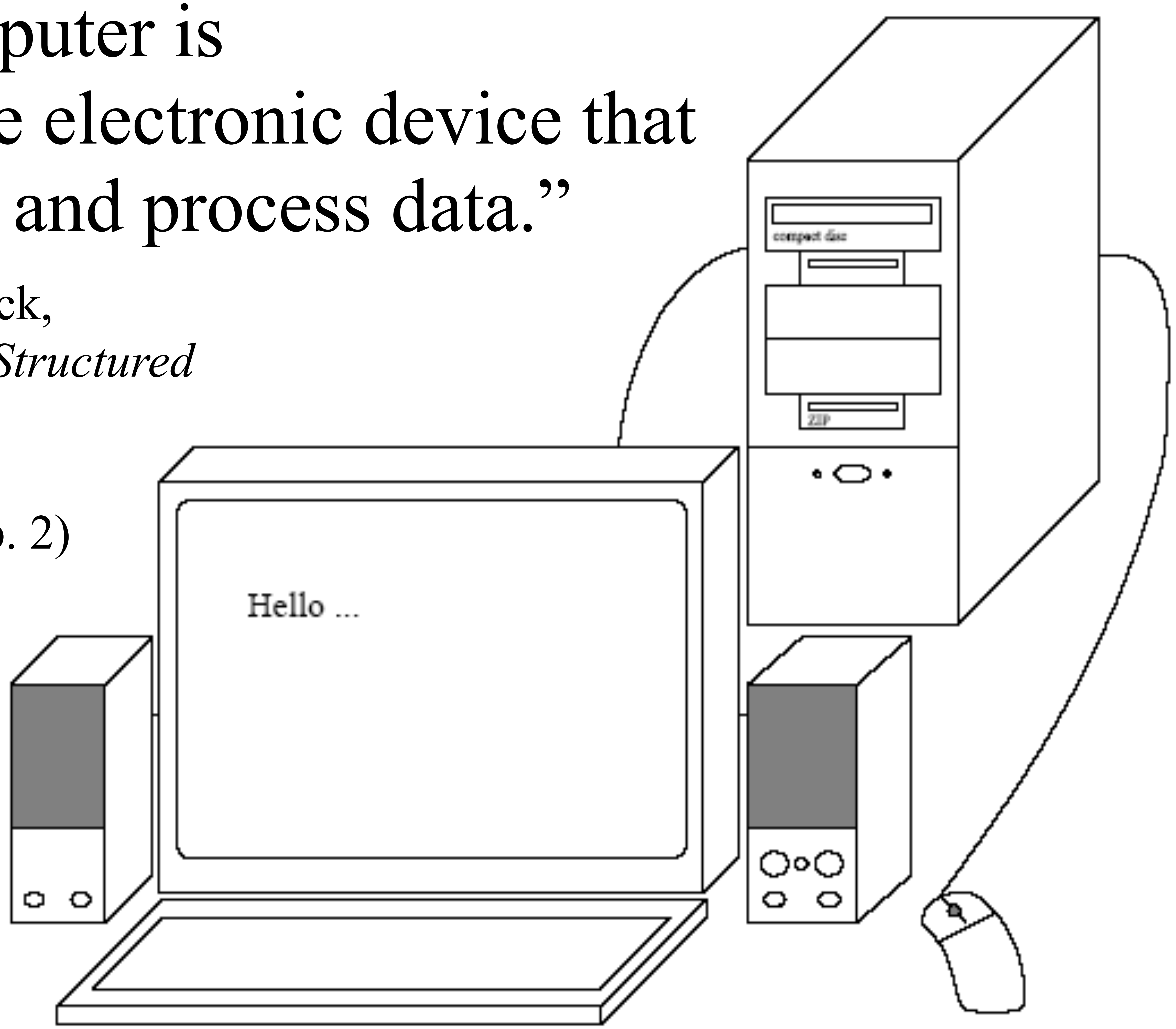
- Central Processor (CPU)
- Random Access Memory (RAM)
- Secondary Memory (HDD/SSD)
- Input devices
    - Mouse
    - Keyboard
    - Net
- Output devices
    - Monitor
    - Net



# What is a Computer?

A computer is
"… [A] programmable electronic device that can store, retrieve and process data."

(N. Dale & D. Orshalick, *Introduction to PASCAL and Structured Design,* D.C. Heath & Co., Lexington MA, 1983, p. 2)



# Today's Laptop

Dell Latitude 5430

https://i.dell.com/is/image/DellContent/content/dam/ss2/product-images/dell-client-products/notebooks/latitude-notebooks/14-5430/pdp/laptop-latitude-14-5430-pdp-hero-504x350.psd?qlt=95&fit=constrain,1&hei=500&fmt=jpg&wid=570&hei=400

- Intel Core i5-1235U 1.30 GHz ("Alder Lake"), 10 cores:
  - 2 performance cores @ 4.4 GHz
  - 8 efficiency cores @ 3.3 GHz
- 32 GB 3200 MHz DDR4 SDRAM
- 512 GB SSD M.2
- 1 Gbps Ethernet Adapter, WiFi

# Moore's "Law"

- "Transistor density on integrated circuits doubles about every two years"
  - Gordon Moore, co founder of Intel
- This means at the same cost you get:
  - twice as many logical components
  - and higher speed (or less power consumption)

# Without Moore's Law: 1 Transistor = $1mm^2$

1959

Planar Transistor $1mm^2$

1

60s

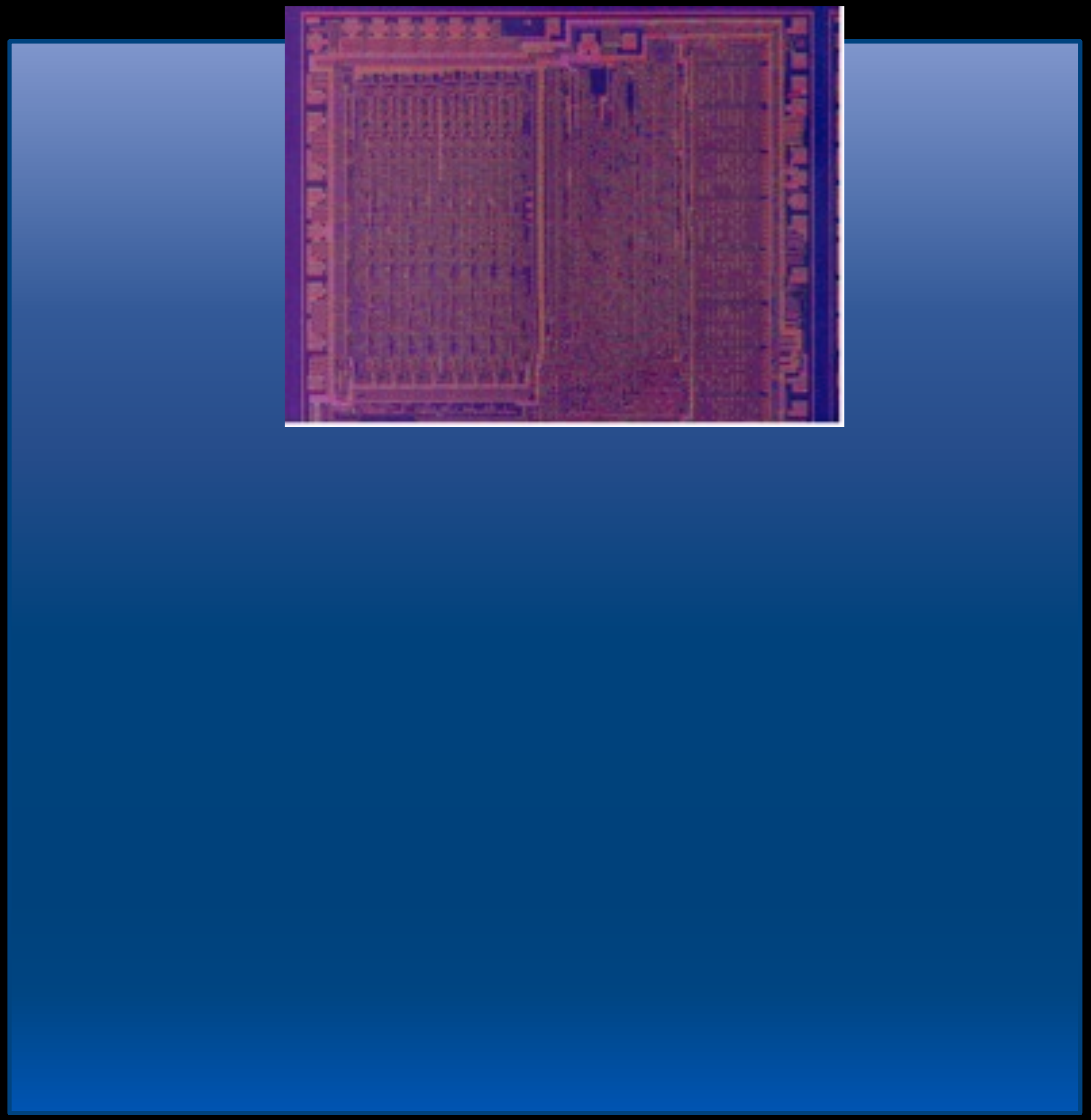
TTL Quad Gate

16

70s

8bit Processor

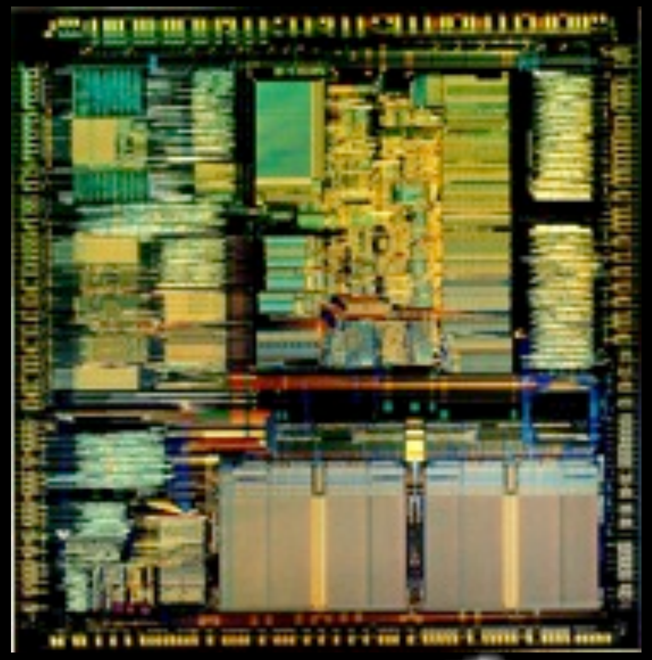
4500



85

32bit Processor

275.000

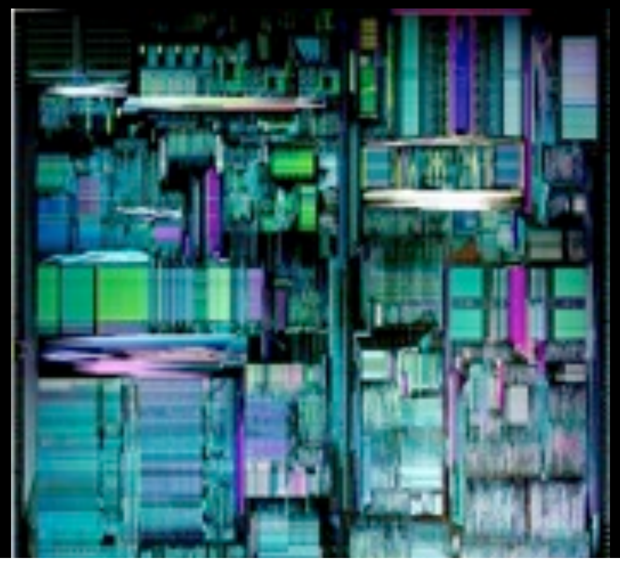


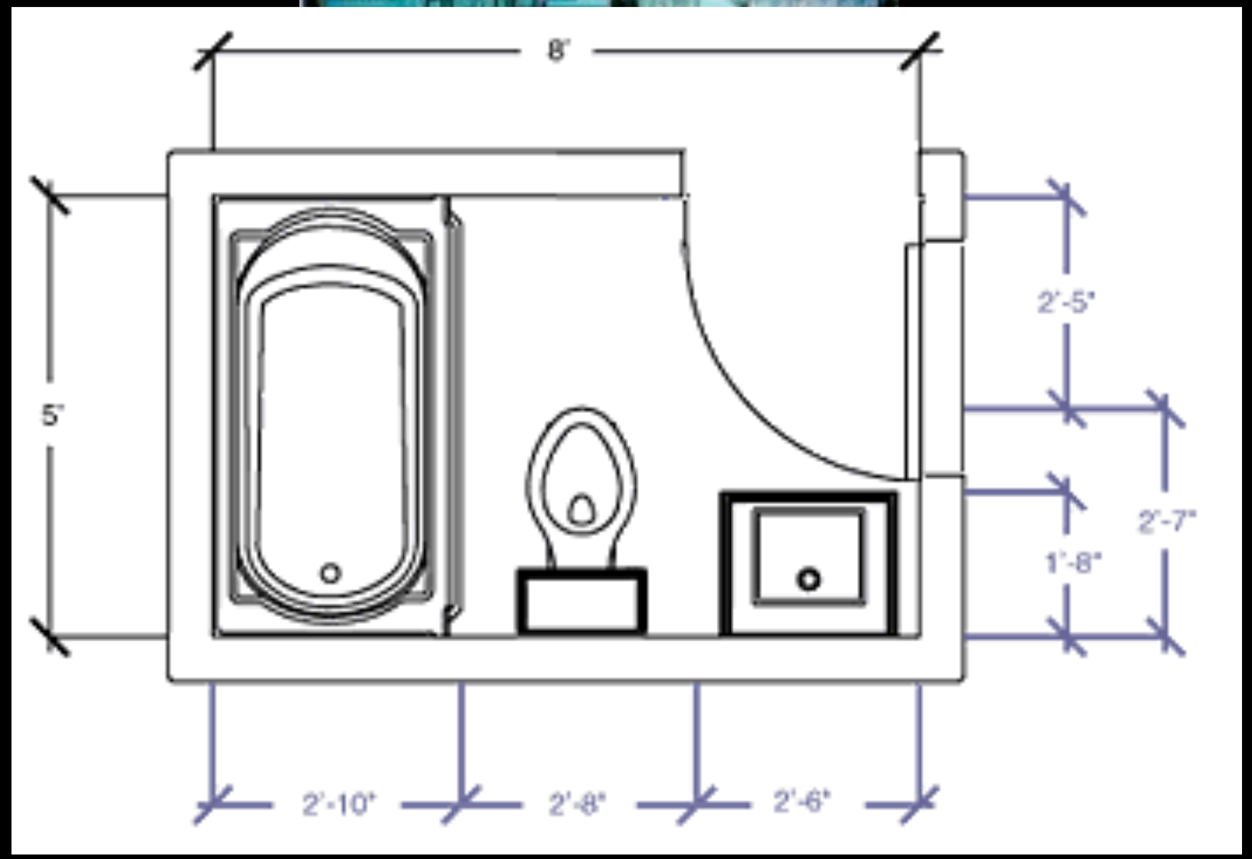
# Without Moore's Law: 1 Transistor = 1mm$^2$

95

32bit Processor

5.500.000

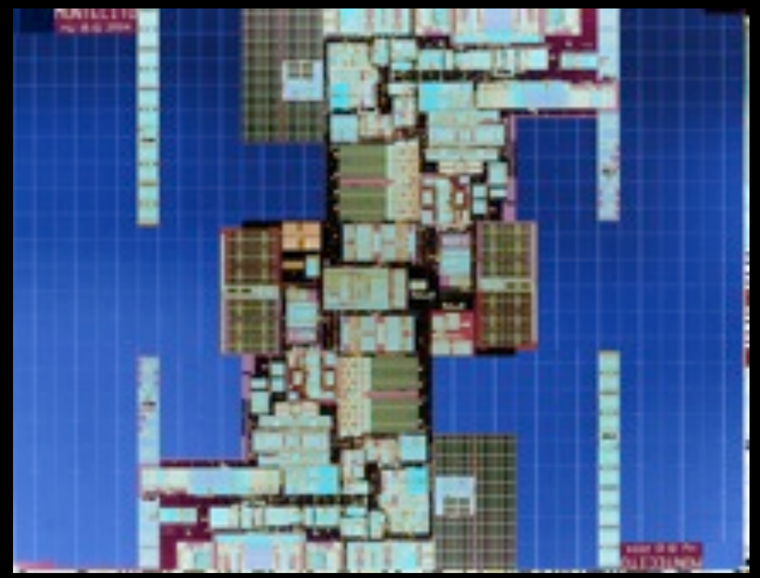





2000s

64bit Processor

221.000.000

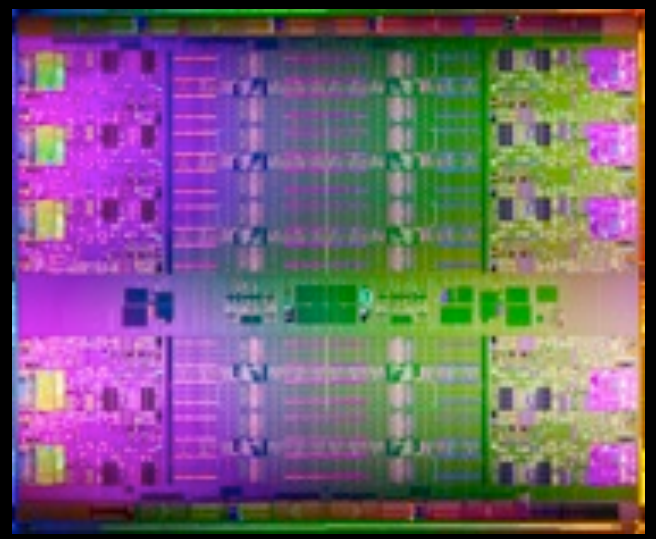


This Room
(- a bit)

2010s

Multi Core Processor

3.000.000.000



Two of those,
plus a bit

# Observations:

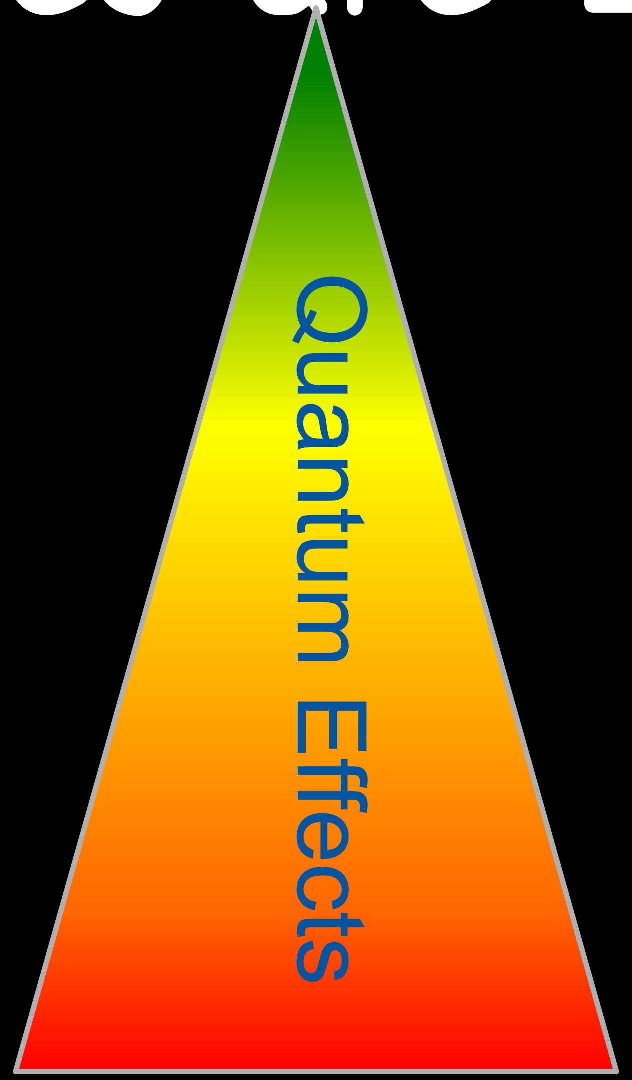
- A 5 Euro music birthday card has more computing power than the Allied Forces during WW-II
- A 2013 smart phone is faster than CERN's computing centre from the mid 90s
- Intuition breaks down → structured approach is needed

# Frequency

- Smaller structures, less electrons need to be pushed around → faster transistors
  - 1970-1990 1→50 MHz (as expected)
  - and then things got interesting....

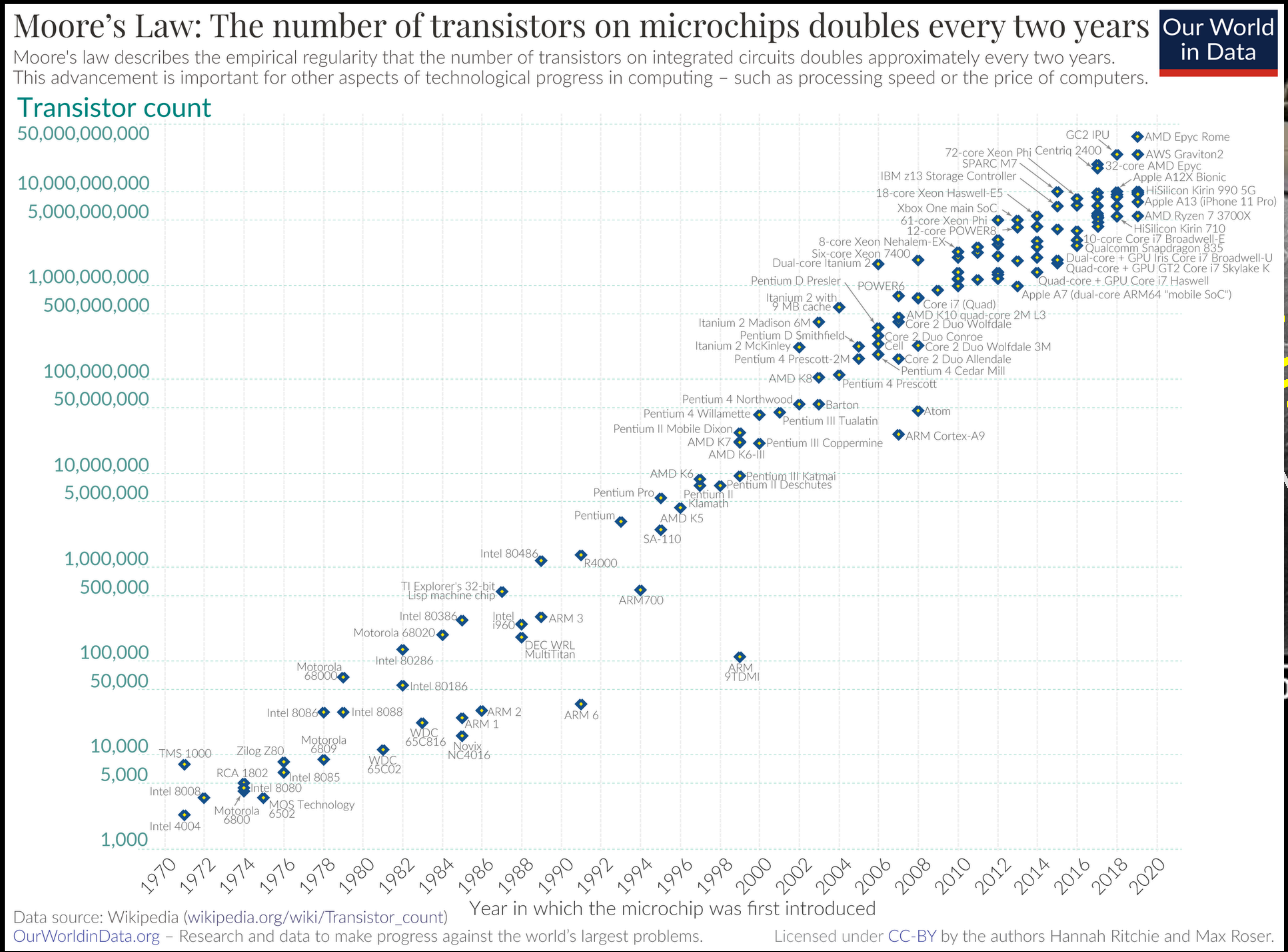
# Moore's Law's Future

- Current structures are 22 nm
  - Intel's roadmap:
    - 14nm 2014
    - 10nm 2016
    - 7nm 2018
    - 5nm 2020
    - 3nm 2022



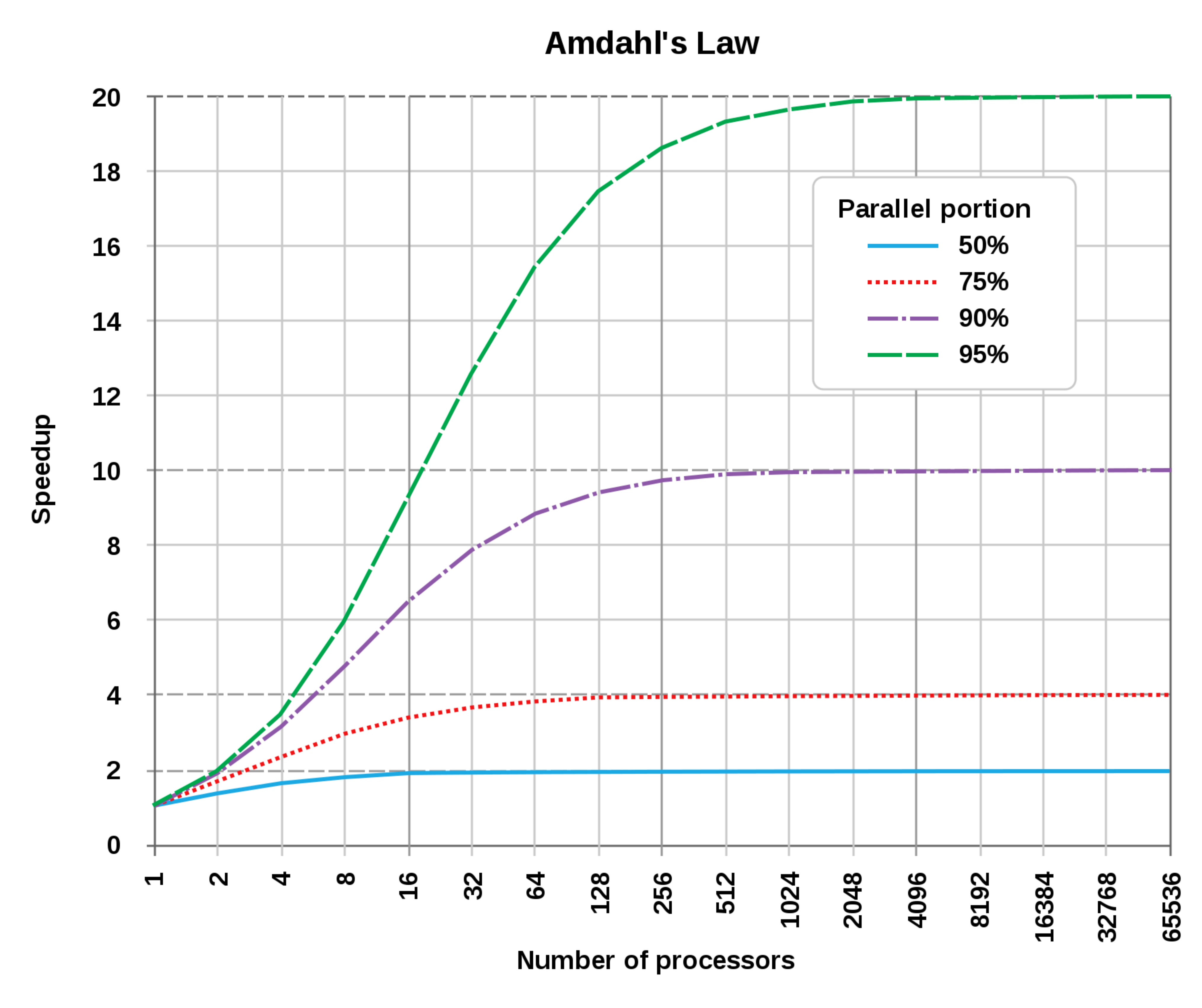

- Smallest transistor 2024
  - 1.5 nm using 1 atom

# Moore's Law's history

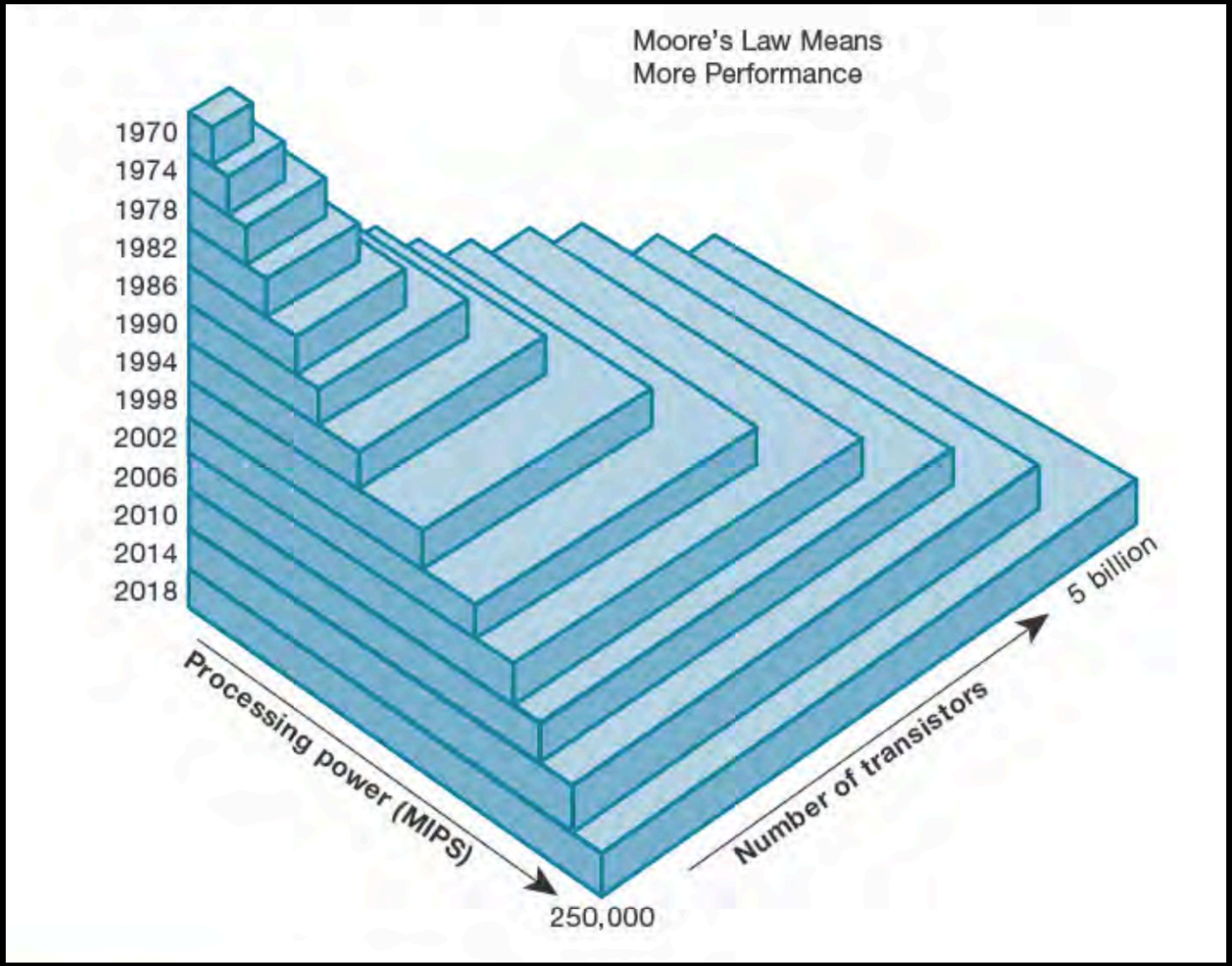



# What to do with 5.000.000.000 Transistors?

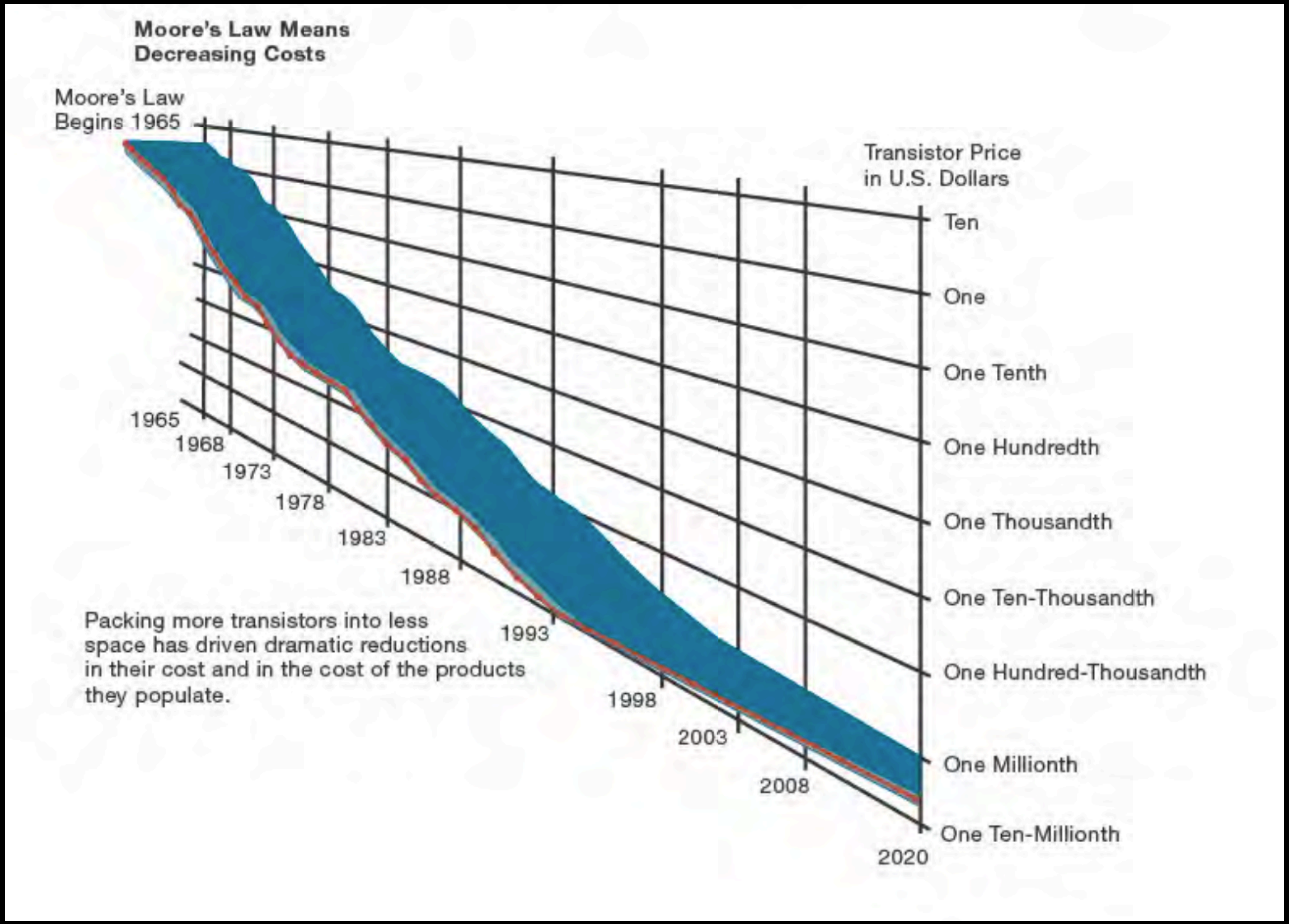
- Two core concepts have been used to speed up computer architectures
  - Pipelining (working like an assembly line)
  - Parallel processing (working concurrently)
- Many additional smart tricks have been invented
  - Caching, branch prediction, etc.....
- Now: CPUs use all tricks combined
  - Difficult to make full use of the hardware
    - Facing sometimes fundamental limits (Amdahl's Law)



# Moore's "Law"



Moore's Law Means
More Performance
1970
1974
1978
1982
1986
1990
1994
1998
2002
2006
2010
2014
2018
Processing power (MIPS)
250,000
Number of transistors
5 billion

Moore's Law Means
Decreasing Costs
Moore's Law
Begins 1965
Transistor Price
in U.S. Dollars
Ten
One
One Tenth
One Hundredth
One Thousandth
One Ten-Thousandth
One Hundred-Thousandth
One Millionth
One Ten-Millionth
1965
1968
1973
1978
1983
1988
1993
1998
2003
2008
2020
Packing more transistors into less
space has driven dramatic reductions
in their cost and in the cost of the products
they populate.

# Implication of Moore's Law

If computing speed and capacity double every 24 months, what are the implications in our lives?

Well, the average undergrad student is – to one significant figure – about 20 years old.

And the average lifespan in the US – to one significant figure – is about 80 years.

So, the average undergrad student has 60 years to go.

So how much will computing speed and capacity increase during the time you have left?

# Double, double, …

60 years / 2 years = 30 doublings

What is $2^{30}$?

Consider the computer on your desktop today, compared to the computer on your desktop the day you die.

How much faster will it be?

Can we possibly predict what the future of computing will enable us to do?

# Computing in 2024

- Moore's law will continue
  - harder and harder, more diversity, complex solutions can gain an edge ( see bitcoin )
  - first quantum computers ( for special purposes, Super Computers )

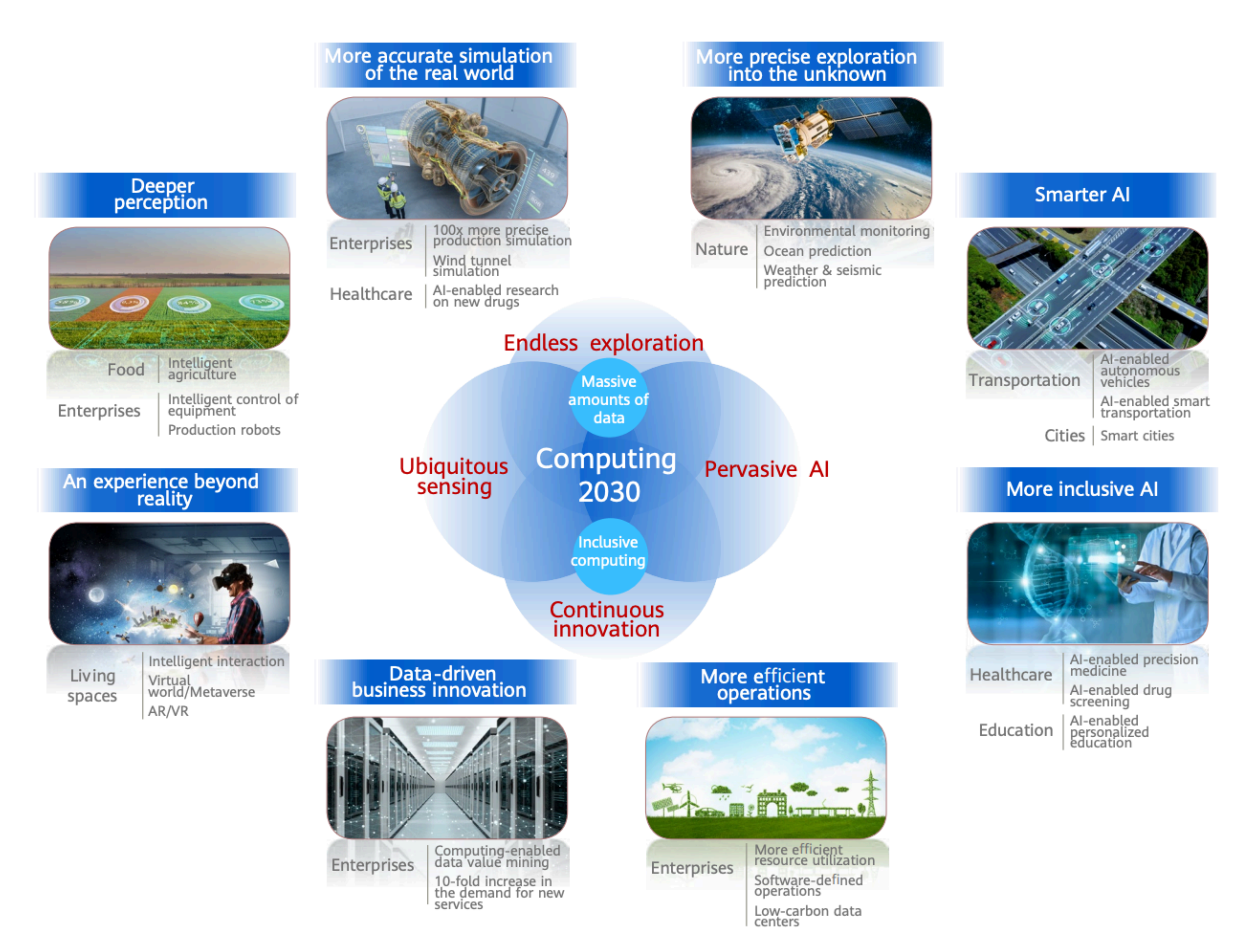
- Computers will change shape
  - most computing and storage is provided by Cloud services
  - network access at high bandwidth from everywhere

- Assistance systems will consume significant computing
  - driving, work, augmented reality, translation, etc.
  - first integration between humans and computer systems ( systems to help disabled people)

- Slight slowdown of progress, software will require huge investments to catch up
  - squeezing the lemon, using larger fraction of the silicon real estate

- 2024 years: CPUs with 30 Billion transistors, running at 6 GHz, disks with 30TByte, FiberOptic networks common

# Computing in 2030



# HARDWARE

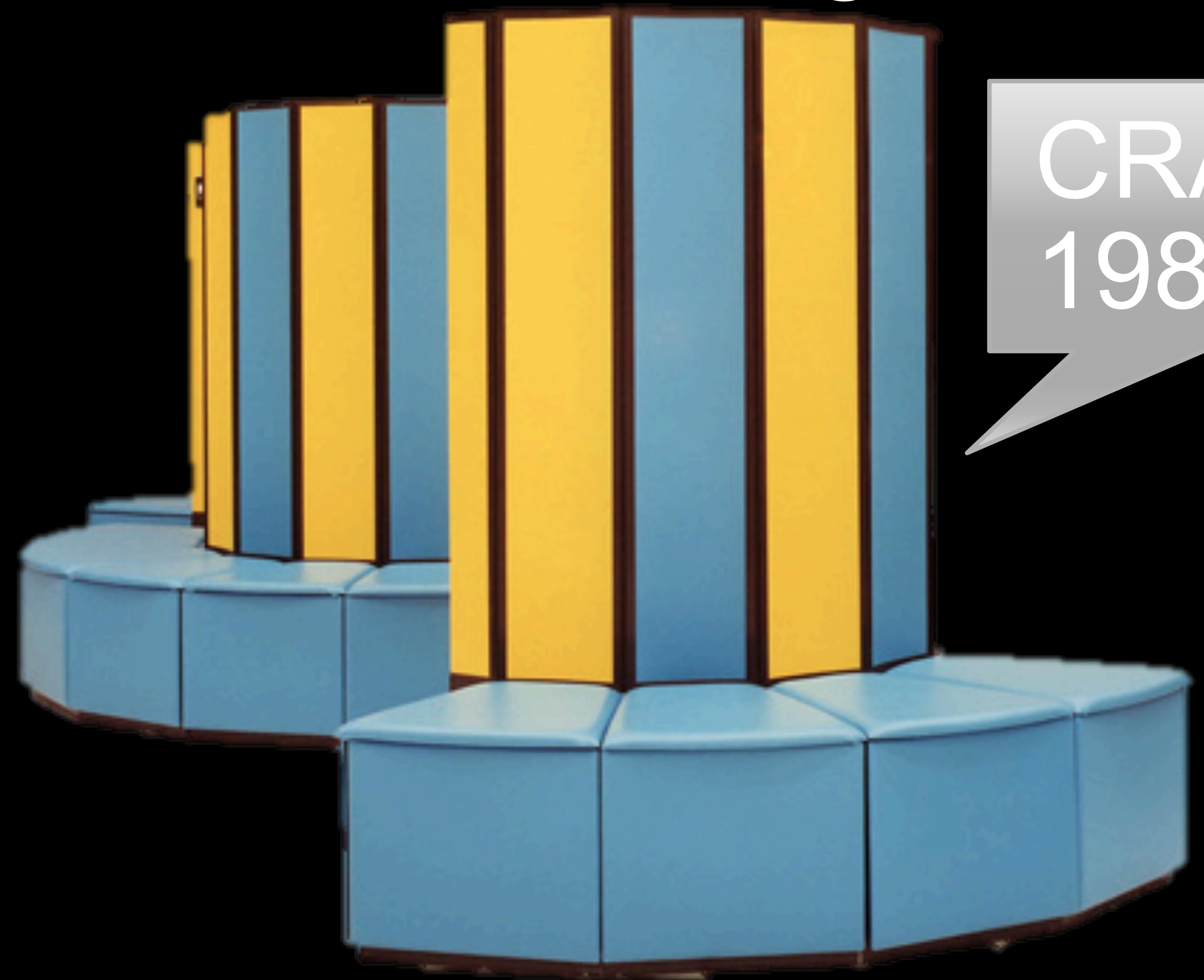
- Supercomputer
- Mainframe
- Server
- Workstation
- Personal computer (PC)
- Mobile device

**TABLE 3.1 Characteristics of Computers Currently Being Used in Organizations**

| Type of Computer | Number of Simultaneous Users | Physical Size | Typical Use | Random Access Memory | Typical Cost (in US$) |
|---|---|---|---|---|---|
| Supercomputer | One to many | Like an automobile to as large as multiple rooms | Scientific research | 5,000+ GB | Up to $200 million |
| Mainframe | 1,000+ | Like a refrigerator | Transaction processing, enterprise-wide applications | Up to 3,000+ GB | Up to $10 million |
| Server | 10,000+ | Like a DVD player and mounted in a rack to fitting on a desktop | Providing websites or access to databases, applications, or files | Up to 512 GB | Up to $50,000 |
| Workstation | Typically one | Fitting on a desktop to the size of a file cabinet | Engineering, medical, graphic design | Up to 512 GB | Up to $10,000 |
| Personal computer | One | Fitting on a desktop | Personal productivity | 512 MB to 32 GB | Up to $5,000 |
| Mobile device | One | Handheld | Personal productivity | 512 MB to 16 GB | Up to $1,400 |

# Super Computers?

- Current Super Computers are build from commodity components (3.120.000 cores)
  - adding a high speed, low latency interconnect
  - adding accelerator cards ( graphic cards, etc. )



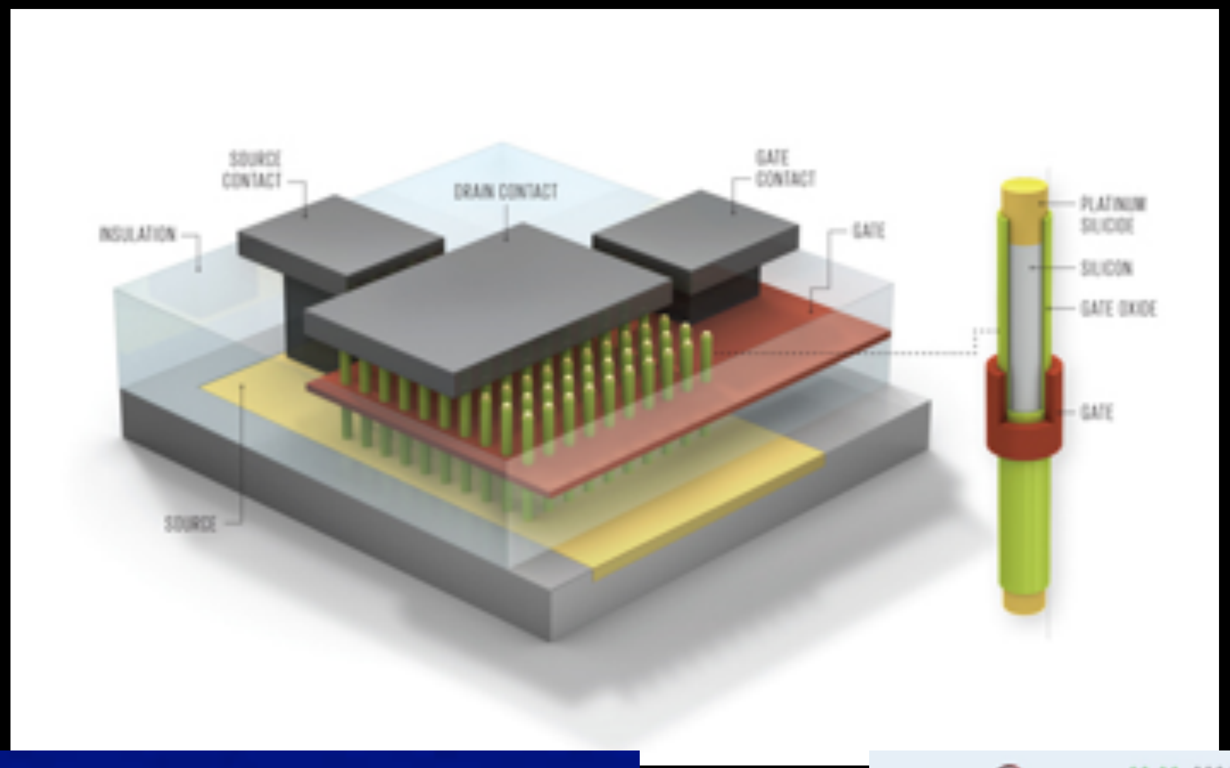

# Pipelining

- Break down operations into small steps
  - which can be done very fast
- Line up the work and finished, complex operations will drop out at the end at a high rate
  - as long as the assembly line is full

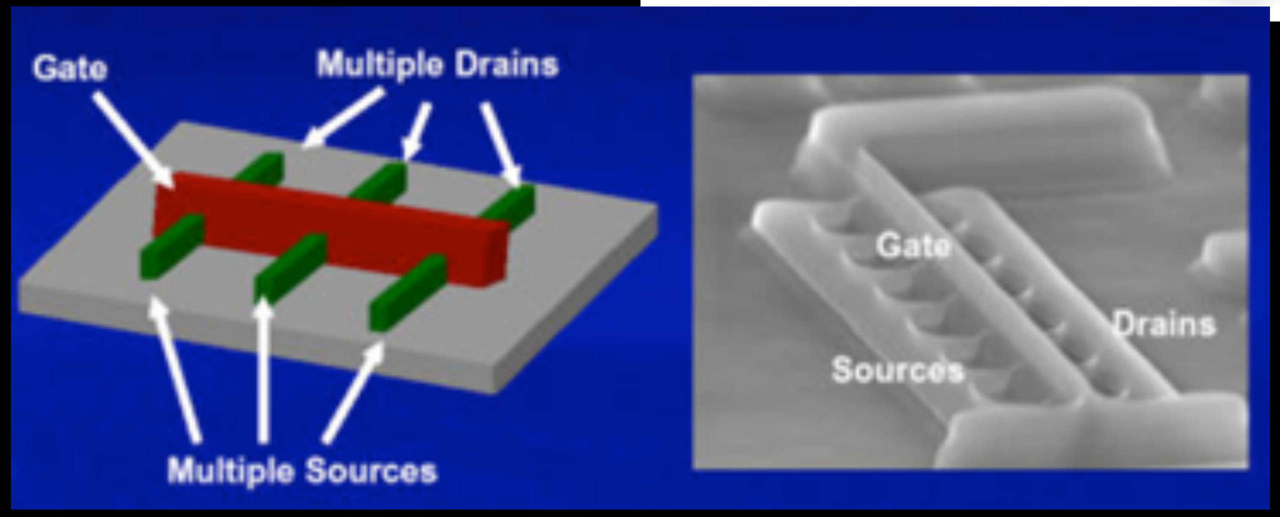
- Limits:
  - Computers are flexible and “react” to results
    - Conditional instructions can break the pipeline
    - In average programs about every 6th instruction
- Works well for many problems, but not for all

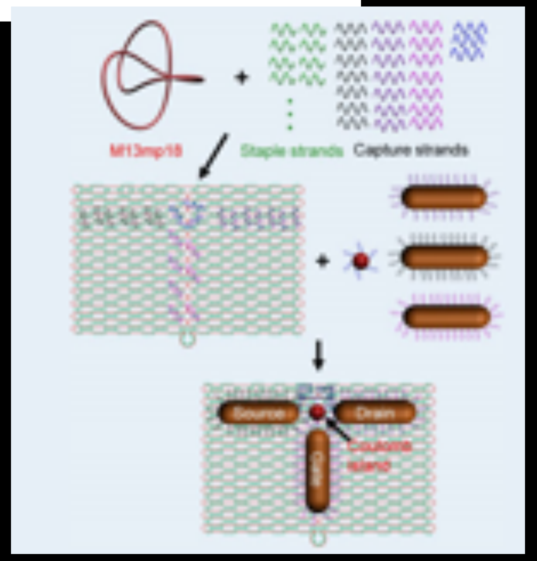
- Pipelines got longer, with diminishing returns

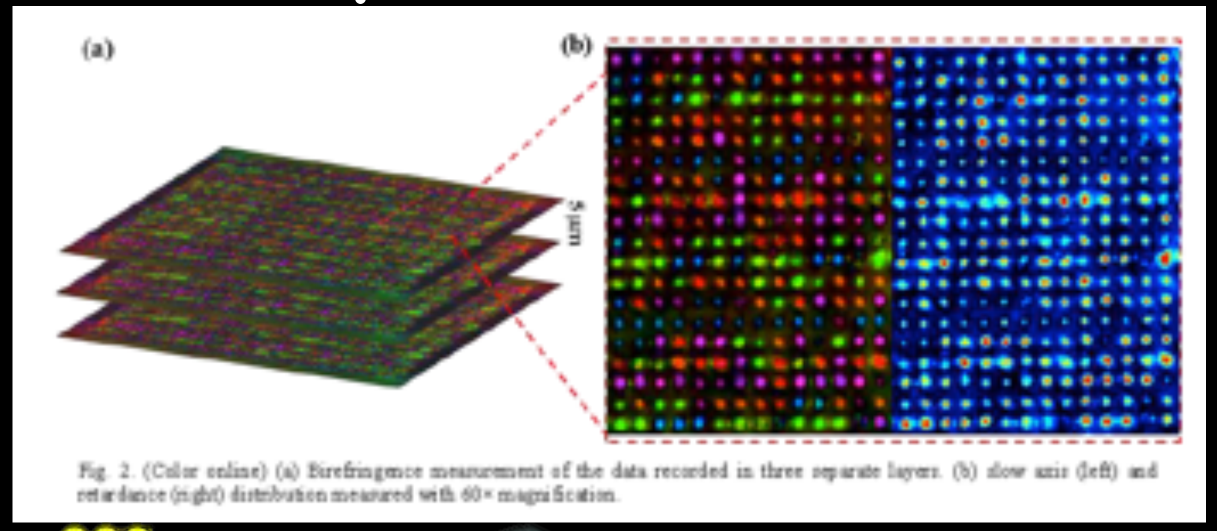
# Alternative Technologies

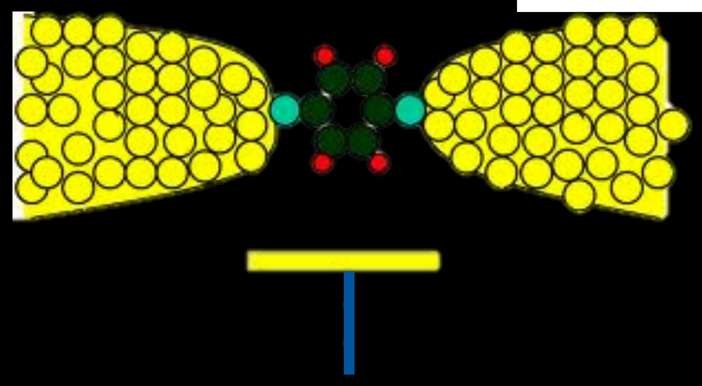


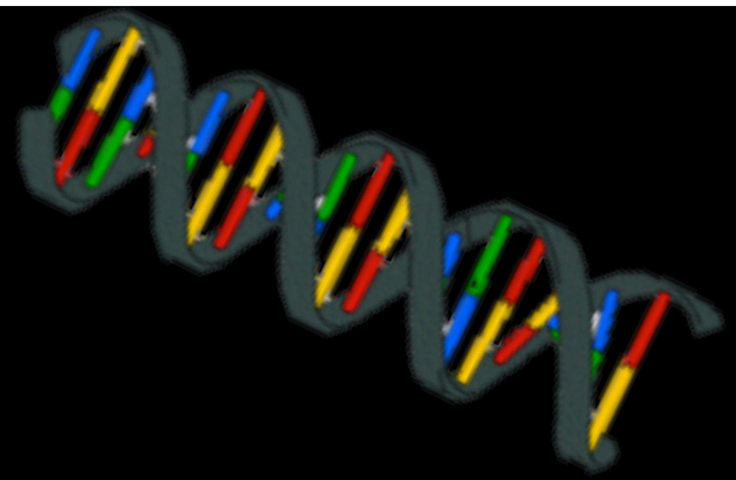
- Nanotubes, Molecular transistors, DNA, Graphen, ......
  - proof of concepts (1-1000 components)
  - many challenges ahead, but promising
- 3D Structures
  - difficult to handle the heat
  - on the surface already widely used
- Memristors, high density, first products soon, but minuscule compared to discs
- Laser nanostructures in glass → 360 TB glass disk
- DNA for storage
  - proof of concept ( $1cm^3$ == 1000 TB)
  - very, very expensive, slow



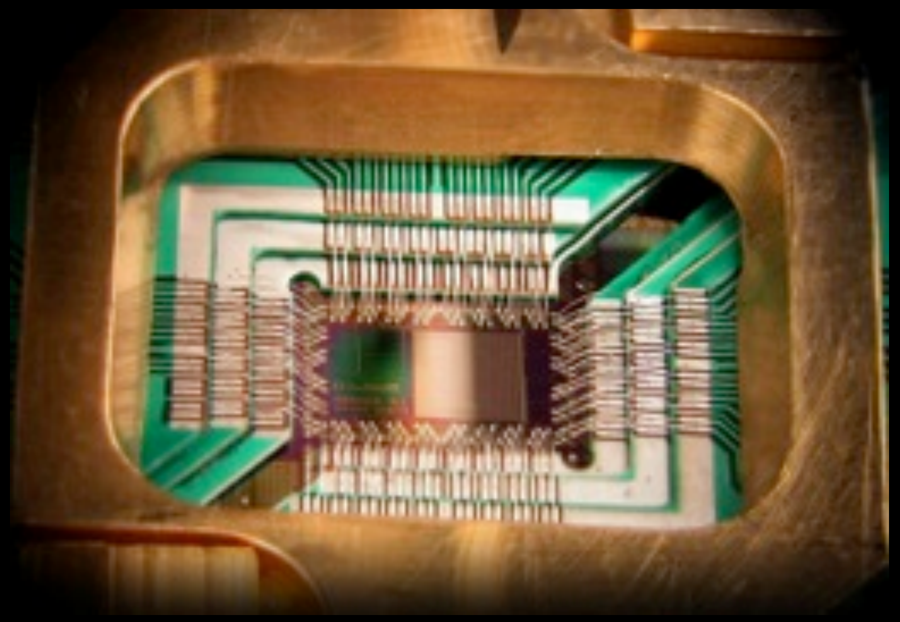









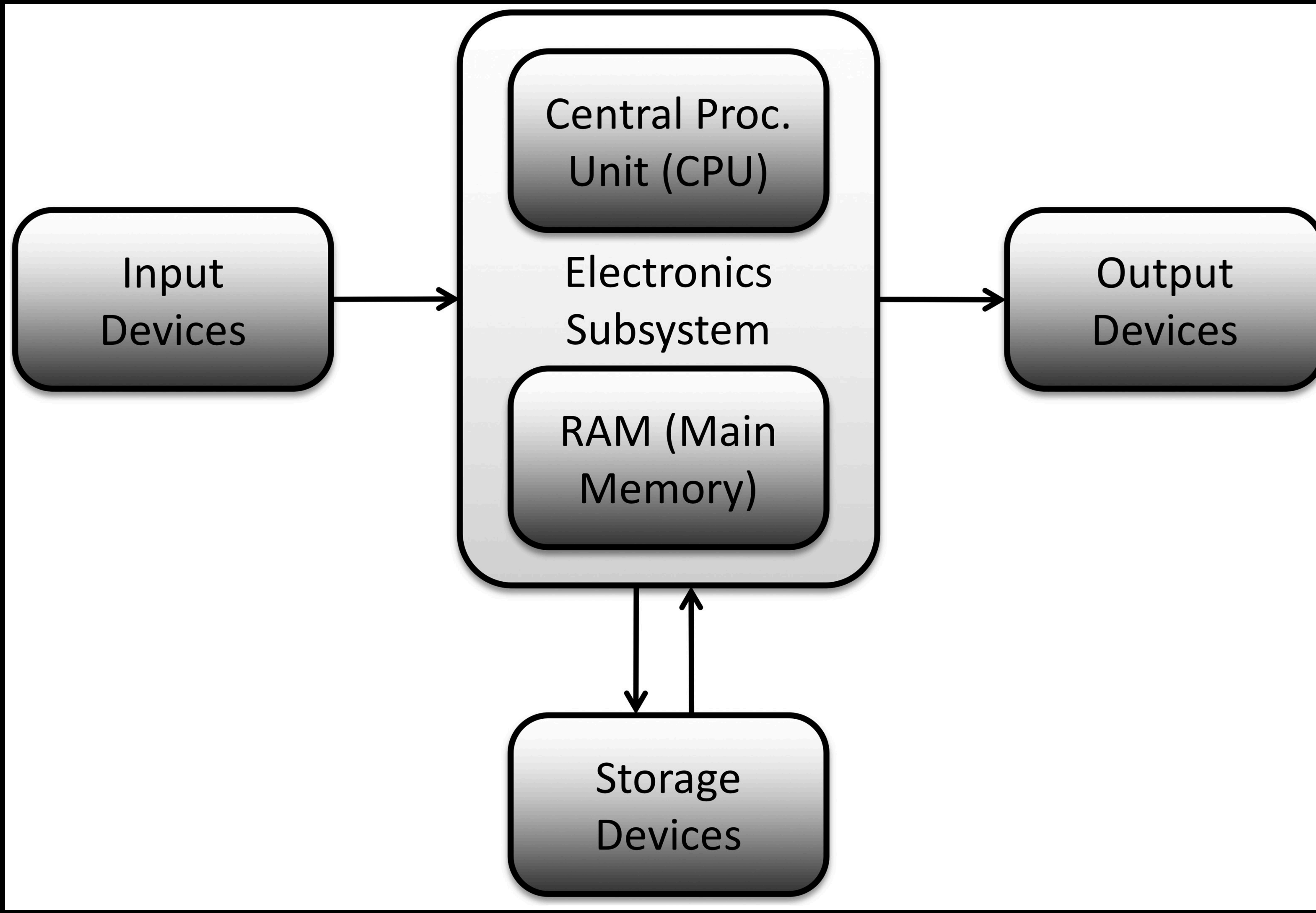
# Quantum Computing



- Two main branches
- Using the phenomenon of quantum state superposition and entanglement to form qubits
  - potential to revolutionize computing
  - systems have to be insulated
  - very early stages, very experimental, many paths are followed
- D-Wave Systems offers a system based on quantum annealing
  - Google and Lockheed are customers
  - system can be used for solving optimization problems
  - most likely to be used as a co-processor

# Hardware - computer

- Central Processor (CPU)
- Random Access Memory (RAM)
- Secondary Memory (HDD/SSD)
- Input devices
    - Mouse
    - Keyboard
    - Net
- Output devices
    - Monitor
    - Net



# Hardware - what to look

- Number of cores (CPU)
- Processor Speed
- Graphics Card (GPU)
- RAM/ROM
- Storage - SSD/HDD
- External devices

| Computer Company, Inc. Model 7000 | |
|---|---|
| *Processor* | Intel® Core™ i9, 8 Cores |
| *Processor Speed* | 4.7GHz |
| *Graphics Processor* | NVIDIA® GeForce RTX™ 2070 |
| *RAM* | 32GB DDR4 RAM at 2666MHz |
| *Disk* | 2TB 7200 RPM SATA |
| *Optical Drive* | Tray Loading DVD-RW Read-Write |
| *Keyboard* | U.S. English |

# CPU Examples

- ***x86***: Intel Celeron/Pentium/Core/Xeon and AMD Ryzen/ Threadripper/EPYC (and related models from smaller manufacturers)

  (Windows, MacOS and Linux PCs; some Android tablets)

  http://en.wikipedia.org/wiki/X86

  Market Share for PCs: Intel 69%, AMD 16%, ARM 15%
  https://www.tomshardware.com/news/amd-and-intel-cpu-market-share-report-recovery-looms-on-the-horizon
- ***ARM*** (99% of smartphones, 15% of laptop/desktop PCs)
  http://en.wikipedia.org/wiki/ARM_processor https://www.xda-developers.com/arm/
- ***IBM POWER10*** (servers) https://en.wikipedia.org/wiki/Power_ISA https://en.wikipedia.org/wiki/Power10



https://substackcdn.com/image/fetch/f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-

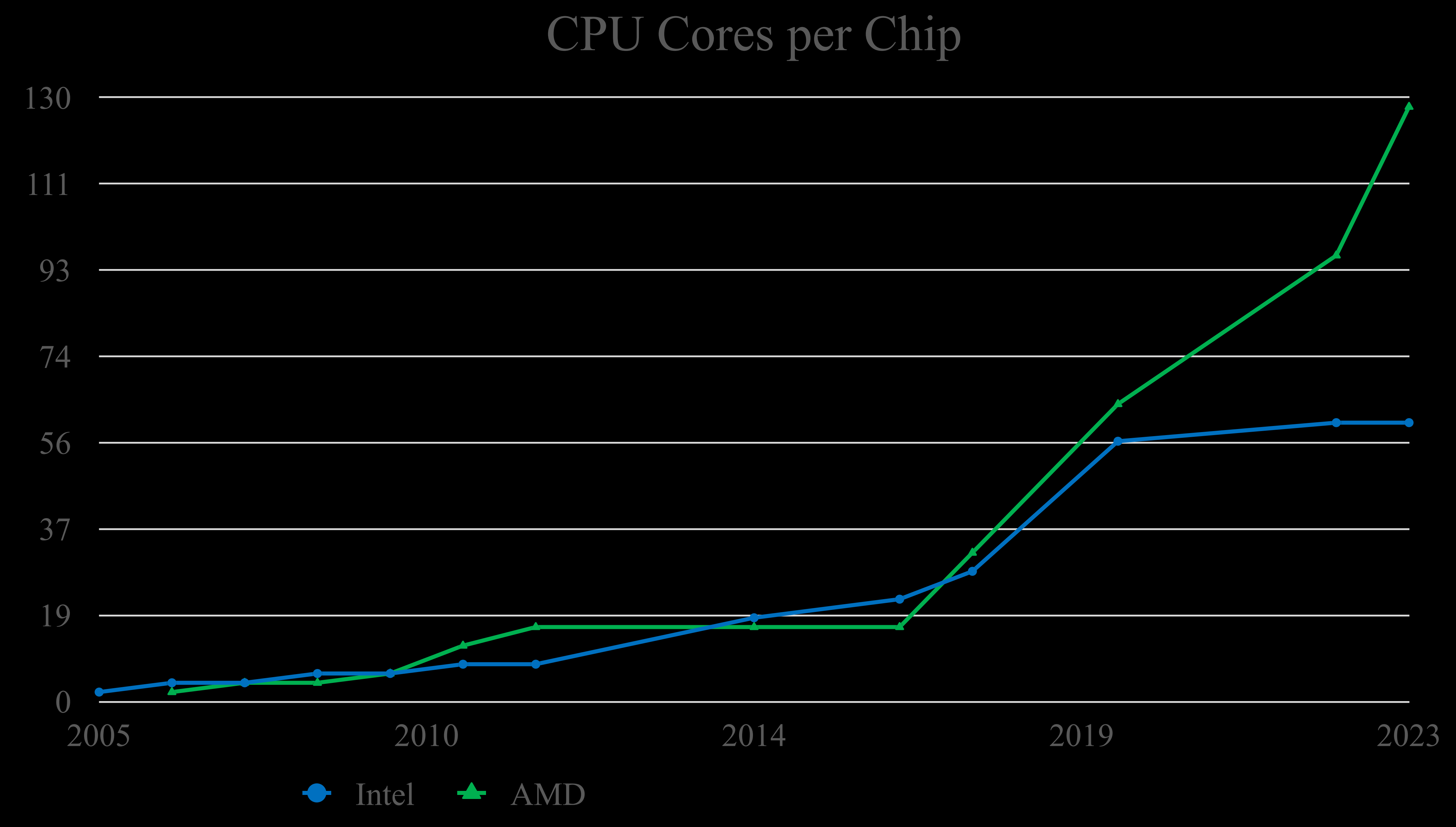
# Multicore

- A ***multicore*** CPU is a chip with multiple, independent "brains," known as ***cores***.
- These multiple cores can run completely separate programs, or they can cooperate together to work simultaneously in parallel on different parts of the same program.
- All of the cores share the same connection to memory – and the same ***bandwidth*** (memory speed).


https://substackcdn.com/image/fetch/f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2Ffe95ad57-d43e-4890-9499-20e2b808e608_6096x4986.jpeg

# Multicore History (x86)

- Single core: November 1971 (Intel 4004)
- Dual core: October 2005 (Intel), March 2006 (AMD)
- Quad core: June 2006 (Intel), Sep 2007 (AMD)
- Hex core: Sep 2008 (Intel), June 2009 (AMD)
- Oct core (Intel & AMD): March 2010
- 12-core (AMD only): March 2010
- 16-core: Nov 2011 (AMD only)
- 18-core: Sep 2014 (Intel only)
- 22-core: March 2016 (Intel only)
- 28-core: July 2017 (Intel only)
- 32-core: June 2017 (AMD only)
- 56-core: Apr 2019 (Intel only)
- 64-core: Aug 2019 (AMD only)
- 96-core: Nov 2022 (AMD only)
- 128-core: June 2023 (AMD only)



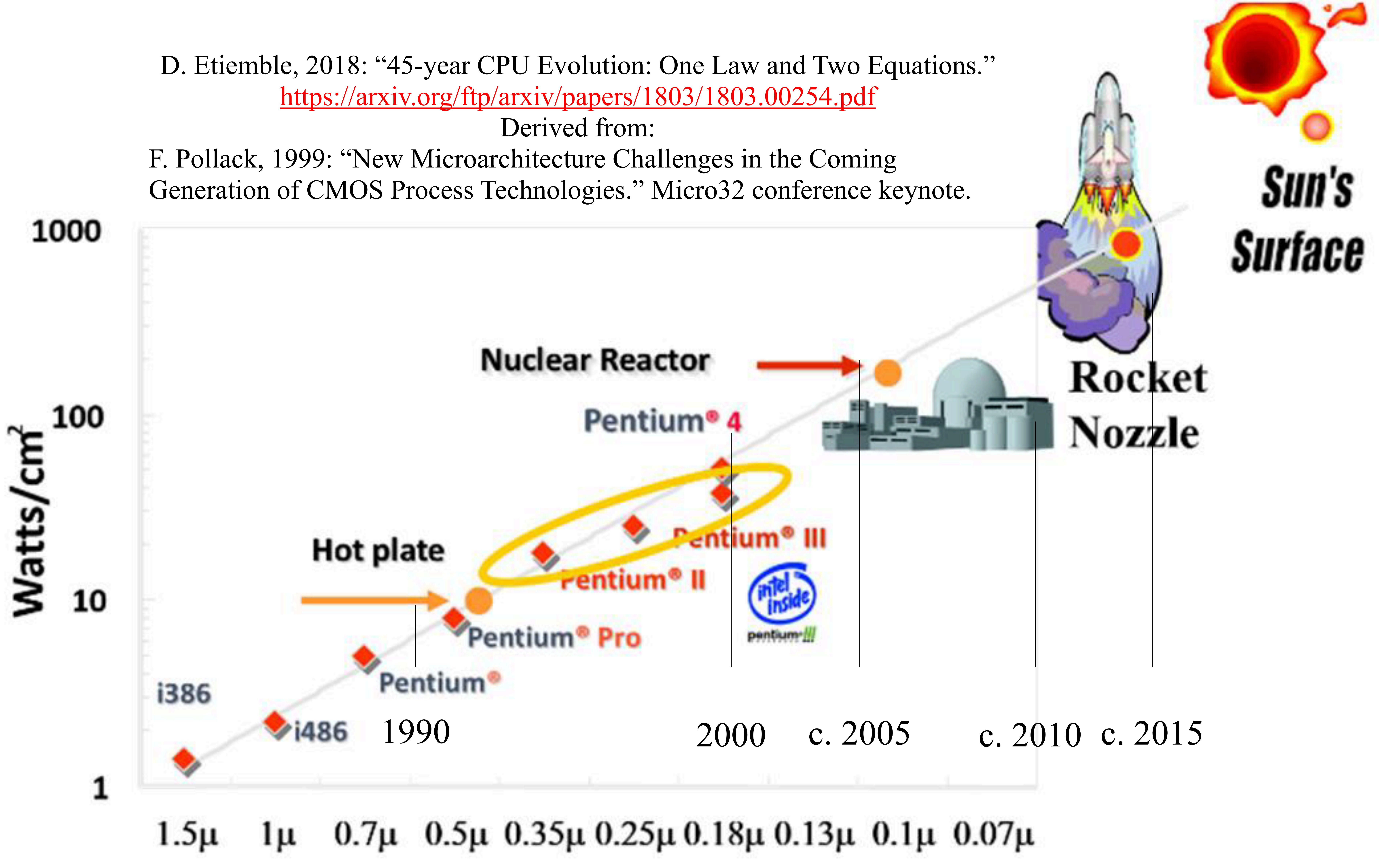

http://www.intel.com/pressroom/kits/quickreffam.htm (dual core, quad core)
http://ark.intel.com/products/family/34348/Intel-Xeon-Processor-7000-Sequence#@Server (hex core)
http://ark.intel.com/ProductCollection.aspx?familyID=594&MarketSegment=SRV (oct core)
http://en.wikipedia.org/wiki/Intel_Nehalem_(microarchitecture) (oct core)
http://en.wikipedia.org/wiki/AMD_Opteron (12-core)
https://en.wikipedia.org/wiki/Broadwell_(microarchitecture) (22-core)
https://en.wikipedia.org/wiki/Skylake_(microarchitecture) (28-core)
https://en.wikipedia.org/wiki/Cascade_Lake_(microarchitecture) (56-core)
https://en.wikipedia.org/wiki/Epyc (32-core, 64-core, 96-core, 128-core)

# Why Multicore? #1

- In the golden olden days (through about 2005), the way to speed up a CPU was to increase its "clock speed."
  - Every CPU has a little crystal inside it that vibrates at a fixed frequency (for example, 1 GHz = 1 billion vibrations per second).
  - Each operation (add, subtract, multiply, divide, etc) requires a specific number of clock ticks to complete.
- But, the power density (watts per square centimeter) of a CPU chip is proportional to the **square** of the clock speed.
- So, continuing to increase the clock speed would have been, quite literally, a dead end, because by now such CPU chips would have already reached the power density of the sun.

# Why Multicore? #2

D. Etiemble, 2018: “45-year CPU Evolution: One Law and Two Equations.”
https://arxiv.org/ftp/arxiv/papers/1803/1803.00254.pdf
Derived from:
F. Pollack, 1999: “New Microarchitecture Challenges in the Coming Generation of CMOS Process Technologies.” Micro32 conference keynote.



# Primary Storage

***Primary storage*** is where data and instructions reside when they're **being used by a program that is currently running**.

- Typically is ***volatile***: The data disappear when the power is turned off.
- Typically comes in two subcategories:
  - Cache
  - Main memory (RAM)

# Main Memory (RAM)

***Main memory*** (RAM) is where data and instructions reside when a **program that is currently running** is going to use them at some point during the run (whether soon or not).

- **Much slower** than cache
  (e.g., less than 1% of CPU speed for RAM, versus
  ~1-100% of CPU speed for cache)
- Therefore, **much cheaper** than cache
  (e.g., ~$0.005/MB for RAM versus $2.34/MB for cache).
  http://www.pricewatch.com/, http://www.ebay.com/,
  http://www.crucial.com/usa/en/compatible-upgrade-for/Dell/latitude-e5540
- Therefore, **much larger** than cache – for example, 1 GB (1024 MB) to 32 TB (~32M MB) for RAM, versus under 1 MB to 1152 MB for cache.
  https://en.wikipedia.org/wiki/Epyc

# RAM versus ROM

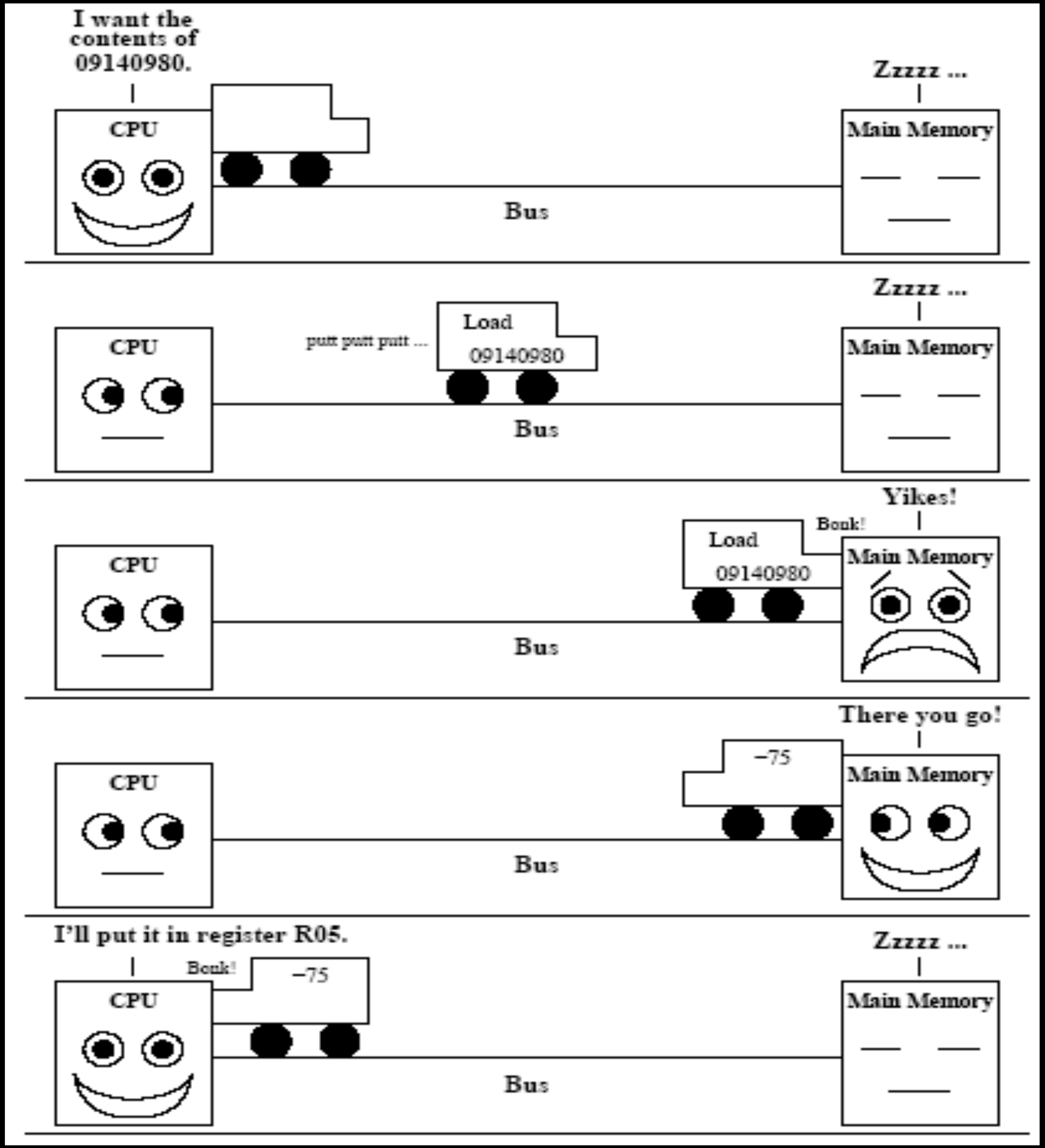
***<u>RAM</u>*: Random Access Memory**

- Memory that the CPU can look at and change arbitrarily (i.e., can load from or store into any location at any time, not just in a sequence).
- We often use the terms **<u>Main Memory</u>**, **<u>Memory</u>** and **<u>RAM</u>** interchangeably.
- Sometimes known as ***<u>core</u>*** memory, named for an older memory technology. (Note that this use of the word "core" is unrelated to "multi-core.")

***<u>ROM</u>*: Read Only Memory**

- Exactly like RAM, except no one can change its values.

# Loading Data from RAM into the CPU

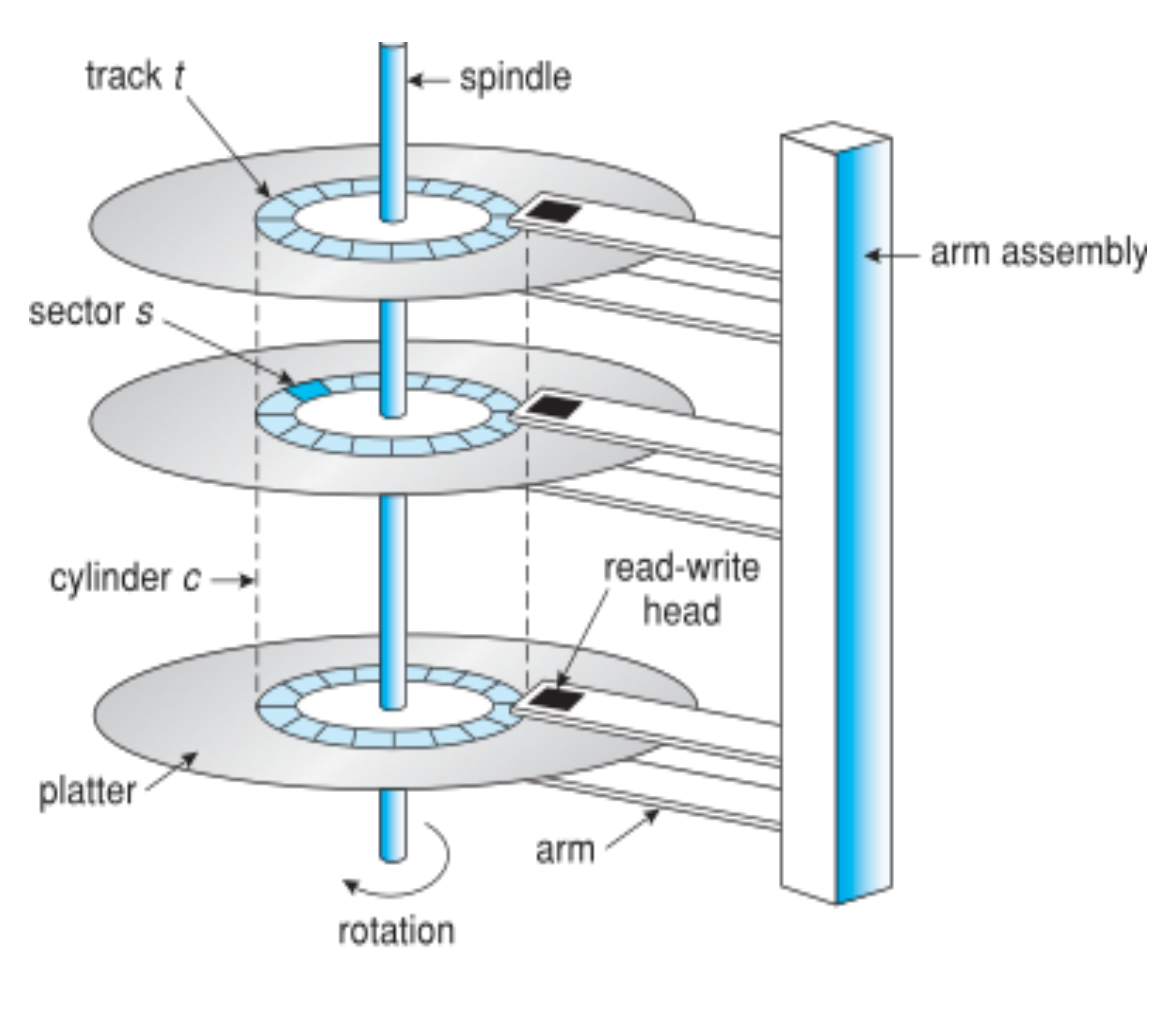


# Strorage and availability

- Uptime refers to the amount of time a system or service is operational and available for use
- Reliability refers to the ability of the system or service to perform its intended function without interruption or failure
- Fault Tolerance refers to the ability of a system or service to continue functioning even in the event of a failure
- Recovery Time  refers to the amount of time it takes for a system or service to recover from a failure
- Scalability refers to the ability of a system or service to handle increasing demands, such as more users or larger amounts of data
- Performance refers to the speed and efficiency of a system or service

| | | Downtime | | |
|---|---|---|---|---|
| Availability % | Friendly Name | per Year | per Month | per Week |
| 90% | one nine | 36.5 days | 72 hours | 16.8 hours |
| 99% | two nines | 3.65 days | 7.2 hours | 1.68 hours |
| 99.5% | -- | 1.83 days | 3.6 hours | 50.4 minutes |
| 99.9% | three nines | 8.76 hours | 43.8 minutes | 10.1 minutes |
| 99.95% | -- | 4.38 hours | 21.56 minutes | 5.04 minutes |
| 99.99% | four nines | 52.56 minutes | 4.32 minutes | 1.01 minutes |
| 99.999% | five nines | 5.26 minutes | 25.9 seconds | 6.05 seconds |
| 99.9999% | six nines | 31.5 seconds | 2.59 seconds | 0.605 seconds |
| 99.99999% | seven nines | 3.15 seconds | 0.259 seconds | 0.0605 seconds |

# Hard Disk - Moving-head Disk Mechanism



# Media Types

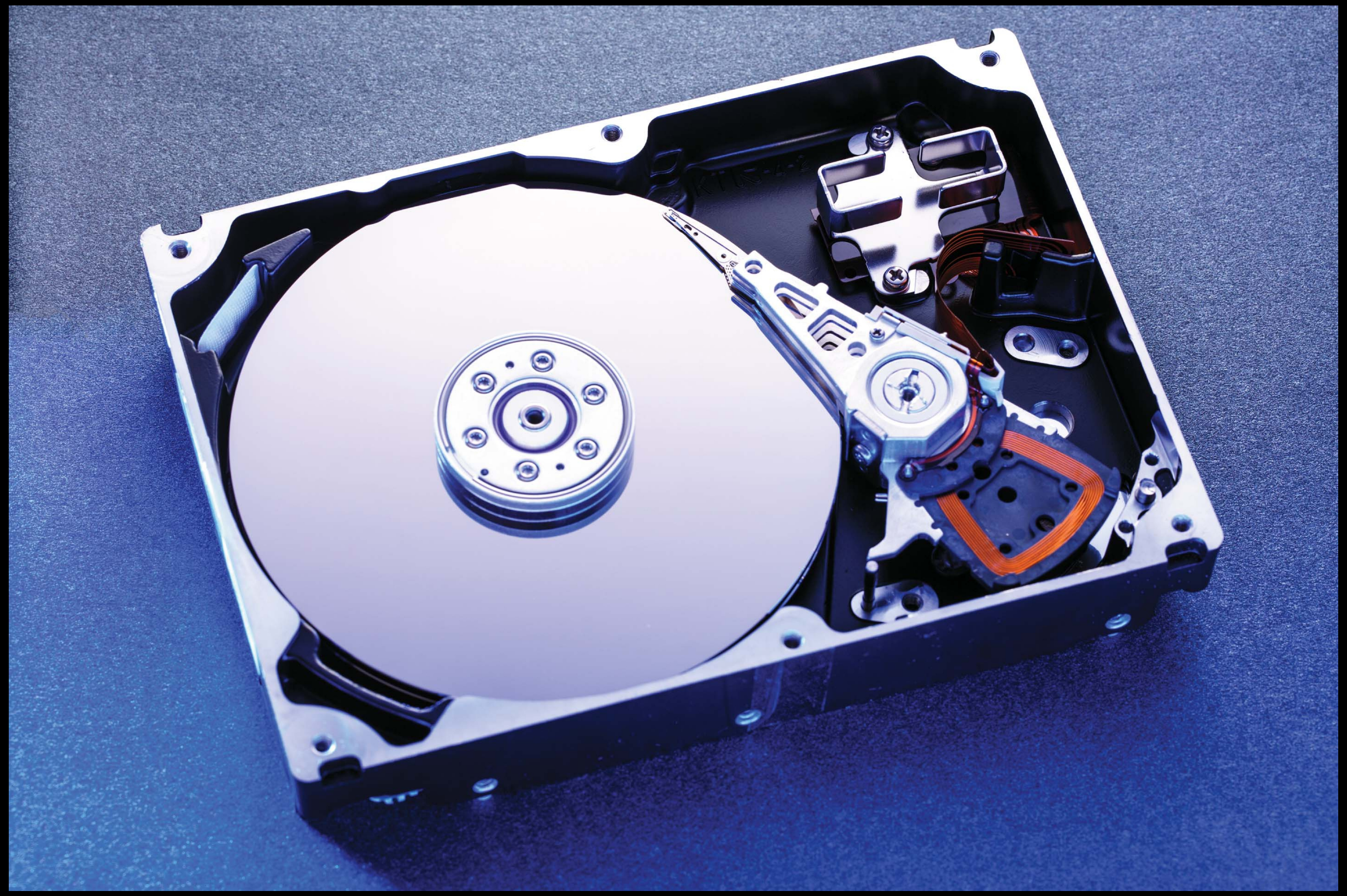
- ***Solid State*** (for example, flash drive)
  - Always can be read
  - Always can be written and rewritten multiple times
  - Contents don't degrade much over time
  - Can't be erased by magnets
- ***Magnetic*** (for example, spinning disk drive)
  - Always can be read
  - Always can be written and rewritten multiple times
  - Contents degrade relatively rapidly over time
  - Can be erased by magnets
- ***Optical*** (for example, DVD)
  - Always can be read
  - Some can be written only once, some can be rewritten multiple times
  - Contents degrade more slowly than magnetic media
  - Can't be erased by magnets
- **Paper**: forget about it!

# Memory speed

| Medium | Speed (MB/sec) | Size (MB) | Media Type | Can write to it? | Port-able? | Pop-ular? | Drive cost ($) | Media cost ($/MB) |
|---|---|---|---|---|---|---|---|---|
| **Cache** | 145,600 | 12 | L1/L2/L3 | Y | N | Req’d | | $2.340000 |
| **RAM** | 21,000 | 33,554,432 | DDR4 | Y | N | Req’d | | $0.003000 |
| **USB 3 Thumb** | 380 | 2,000,000 | Solid | Y | Y | Y | | $0.000300 |
| **Hard Disk** | 100 | 20,000,000 | Mag | Y | N | Y | | $0.000014 |
| **Blu-ray** | 72 | 50,000 | Opt | Y | Y | N | $70 | $0.000038 |
| **DVD±RW** | 32 | 4,700 | Opt | Y | Y | N | $13 | $0.000100 |
| **CD-RW** | 7.8 | 700 | Opt | Y | Y | N | $12 | $0.000900 |
| **Mag tape** | 400 | 18,000,000 | Mag | Y | Y | N | $4455 | $0.000004 |
| **Floppy** | 0.03 | 1.44 | Mag | Y | Y | N | $17 | $0.900000 |
| **Cassette** | << 1 | << 1 | Mag | Y | Y | | Historical | |
| **Paper tape** | << 1 | << 1 | Paper | Y | Y | | Historical | |
| **Punch card** | << 1 | << 1 | Paper | Y | Y | | Historical | |

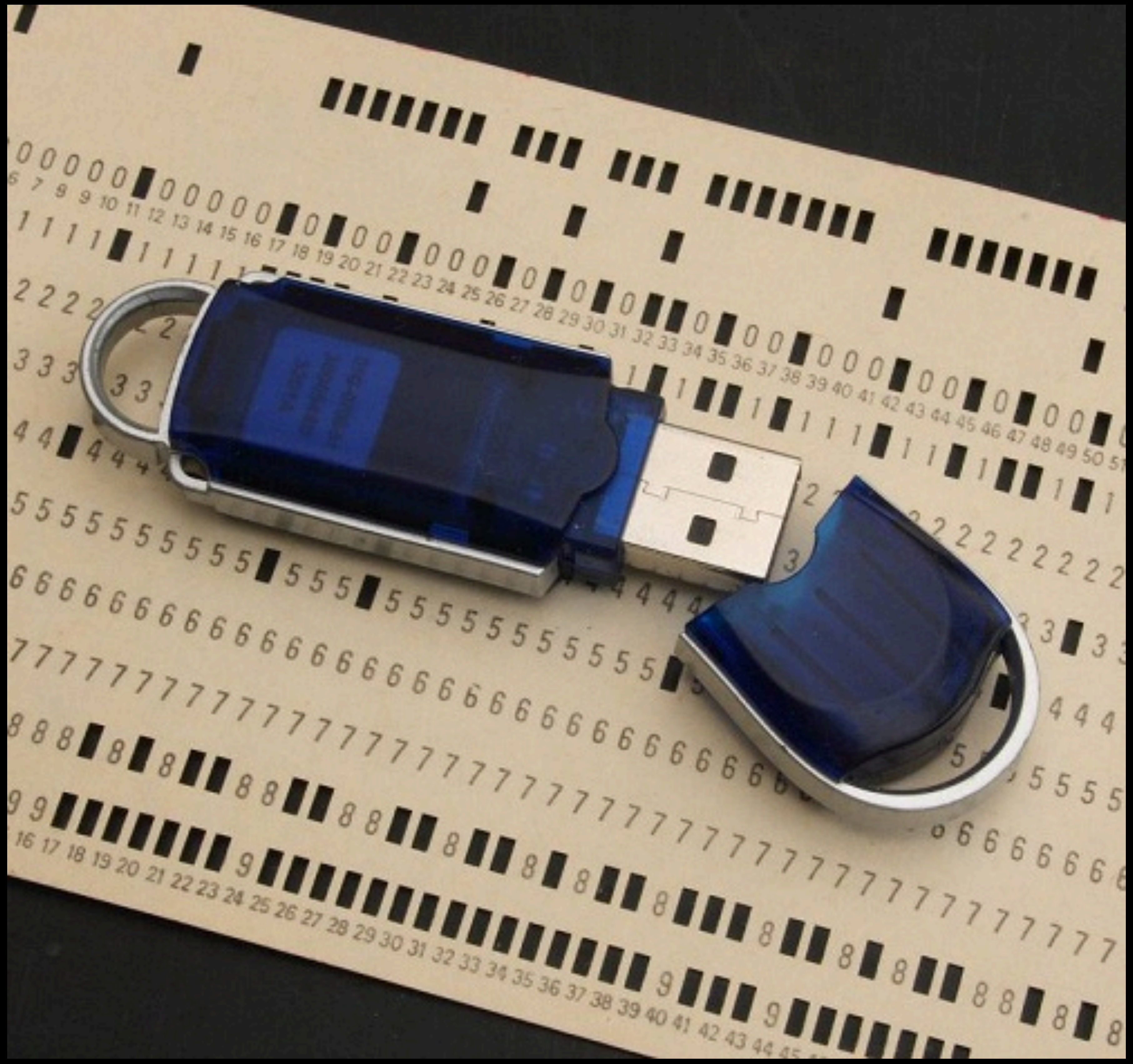
# Hard Disk Drives

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



# Nonvolatile Memory Devices

- If disk-drive like, then called solid-state disks (SSDs)
- Other forms include USB drives (thumb drive, flash drive), DRAM disk replacements, surface-mounted on motherboards, and main storage in devices like smartphones
- Can be more reliable than HDDs
- More expensive per MB
- Maybe have shorter life span – need careful management
- 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



# Nonvolatile Memory Devices

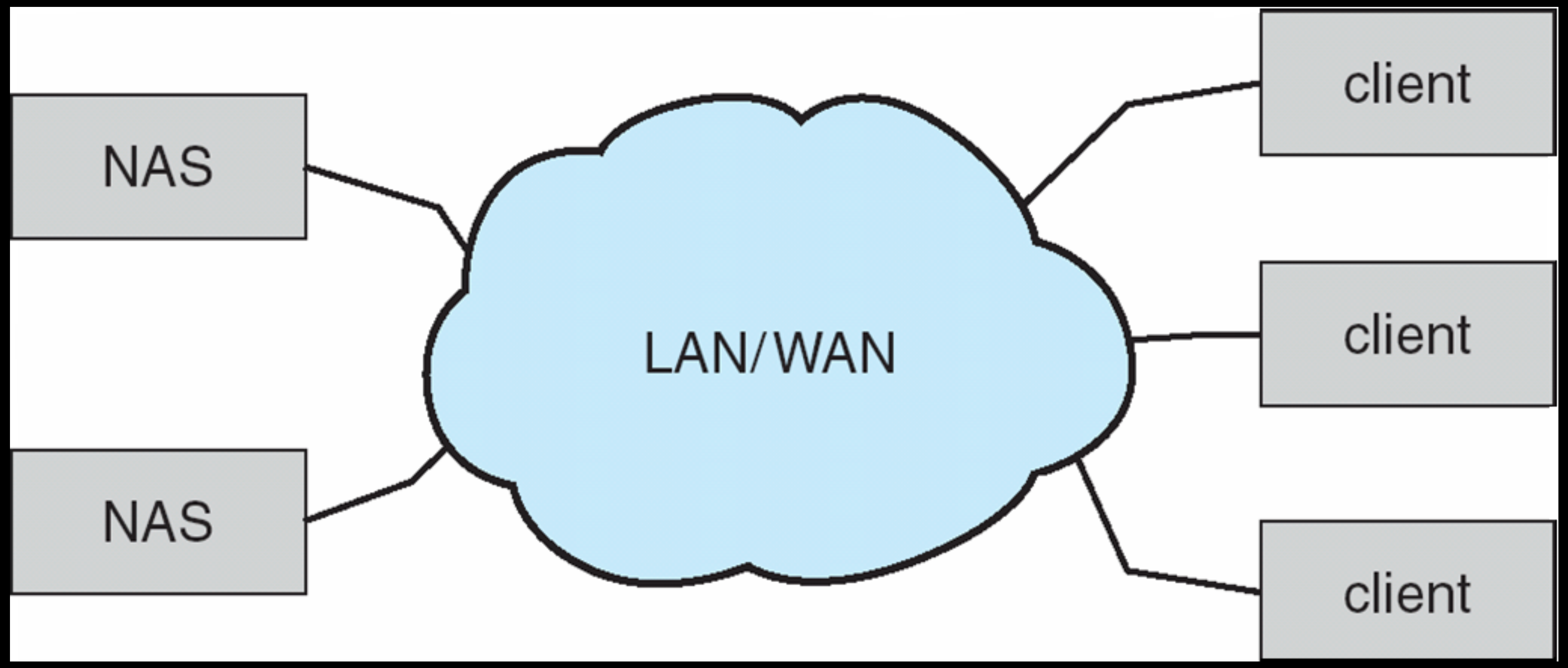
- Have characteristics that present challenges
- Read and written in “page” increments (think sector) but can’t overwrite in place
  - Must first be erased, and erases happen in larger ”block” increments
  - Can only be erased a limited number of times before worn out – ~ 100,000
  - Life span measured in drive writes per day (DWPD)
    - A 1TB NAND drive with rating of 5DWPD is expected to have 5TB per day written within warrantee period without failing

# HDD Scheduling

- The operating system is responsible for using hardware efficiently — for the disk drives, this means having a fast access time and disk bandwidth
- Minimize seek time
- Seek time ≈ 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

# 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**
  - Remotely attaching to devices (blocks)

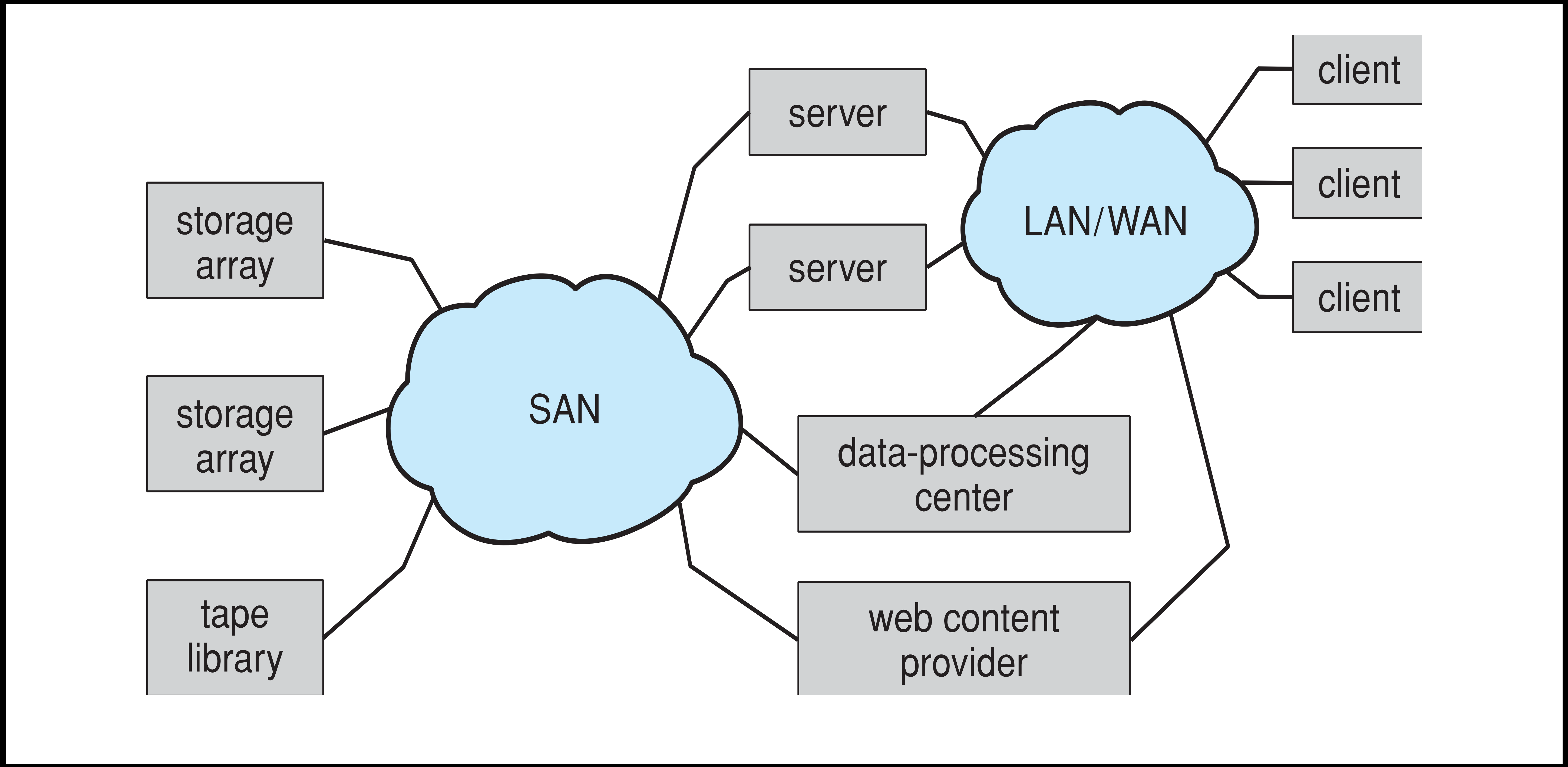


# Storage Array

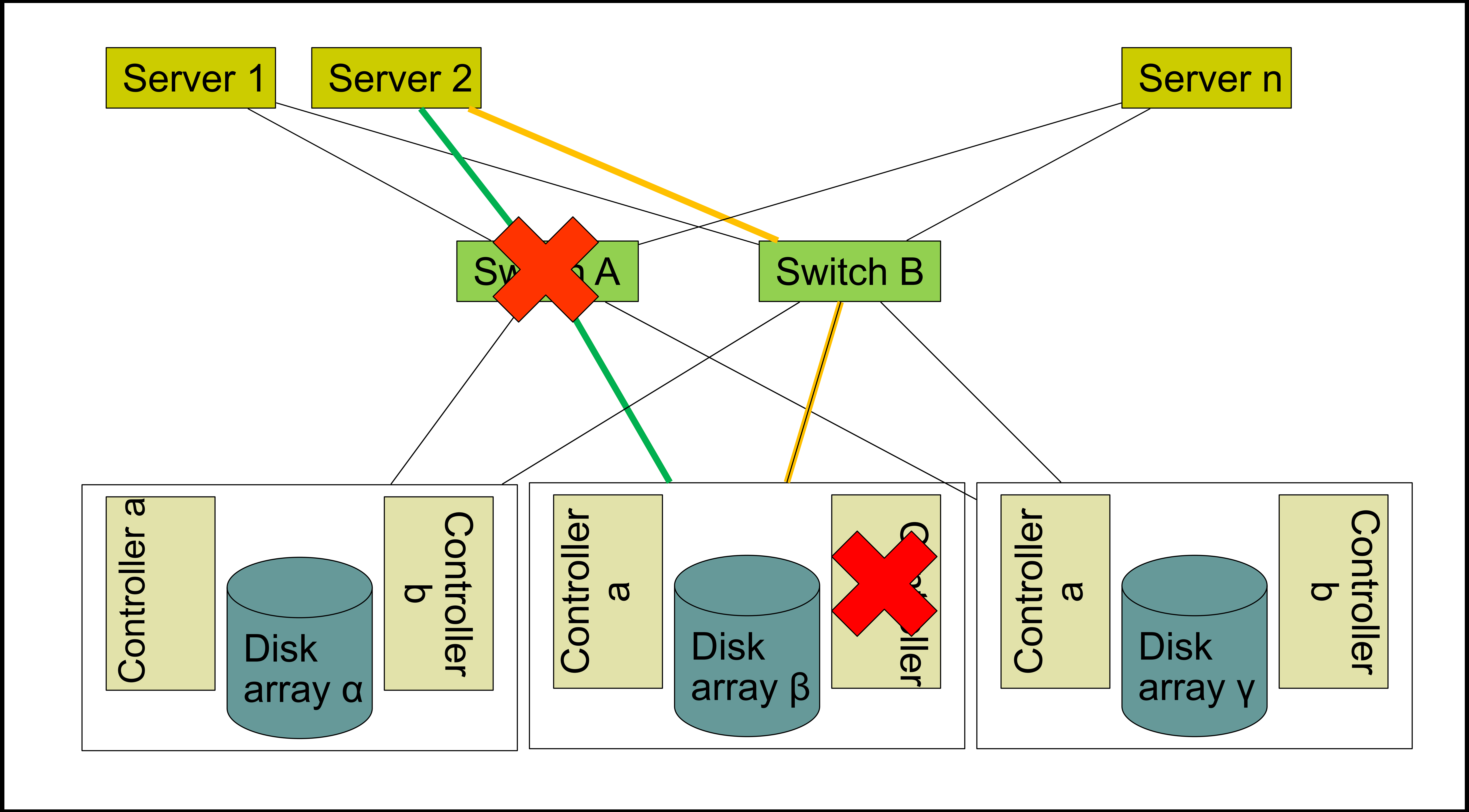
- Can just attach disks, or arrays of disks
- Avoids the NAS drawback of using network bandwidth
- Storage Array has controller(s), provides features to attached host(s)
  - Ports to connect hosts to array
  - Memory, controlling software (sometimes NVRAM, etc)
  - A few to thousands of disks
  - RAID,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

# Storage Area Network

- Common in large storage environments
- Multiple hosts attached to multiple storage arrays – flexible



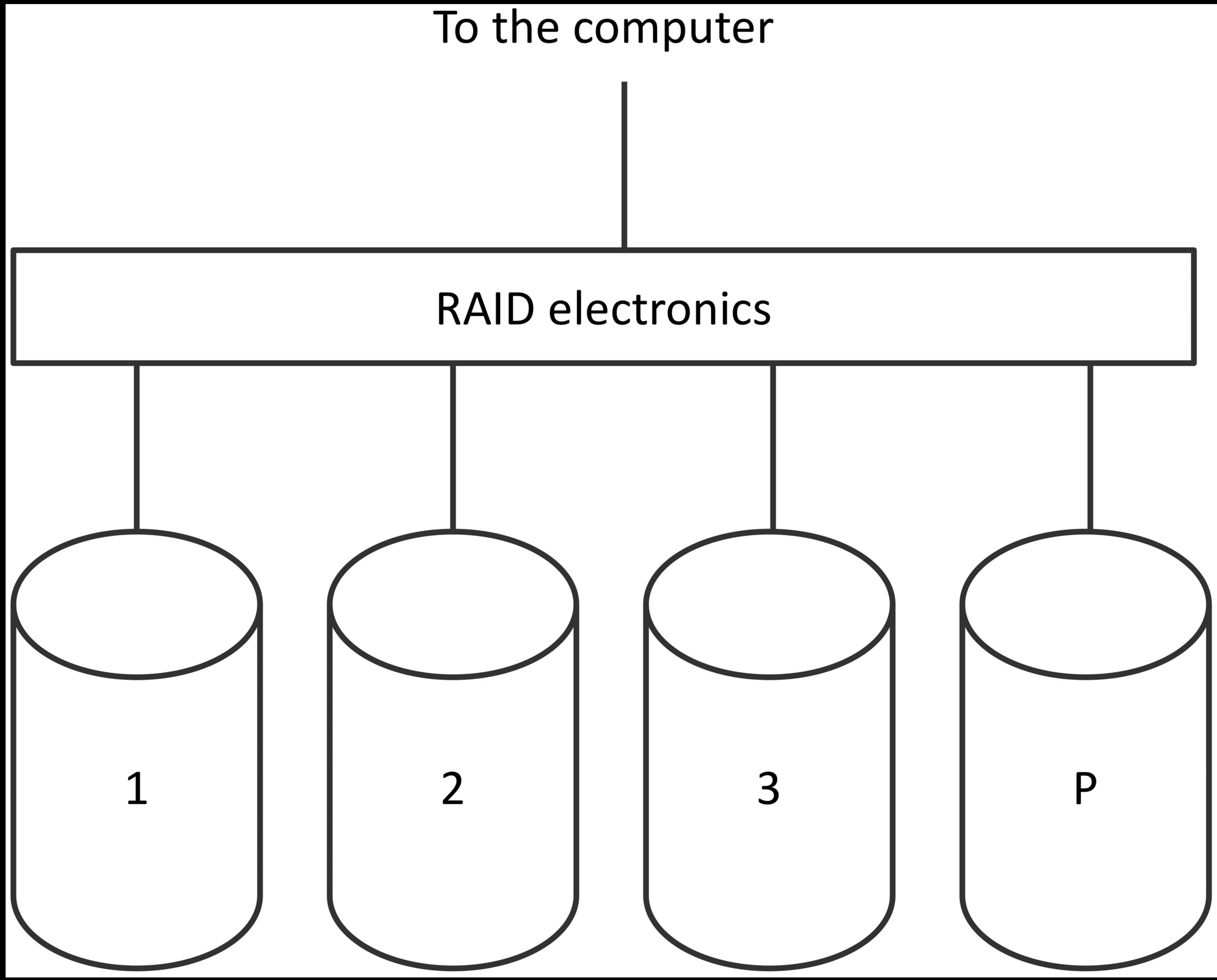
# SAN



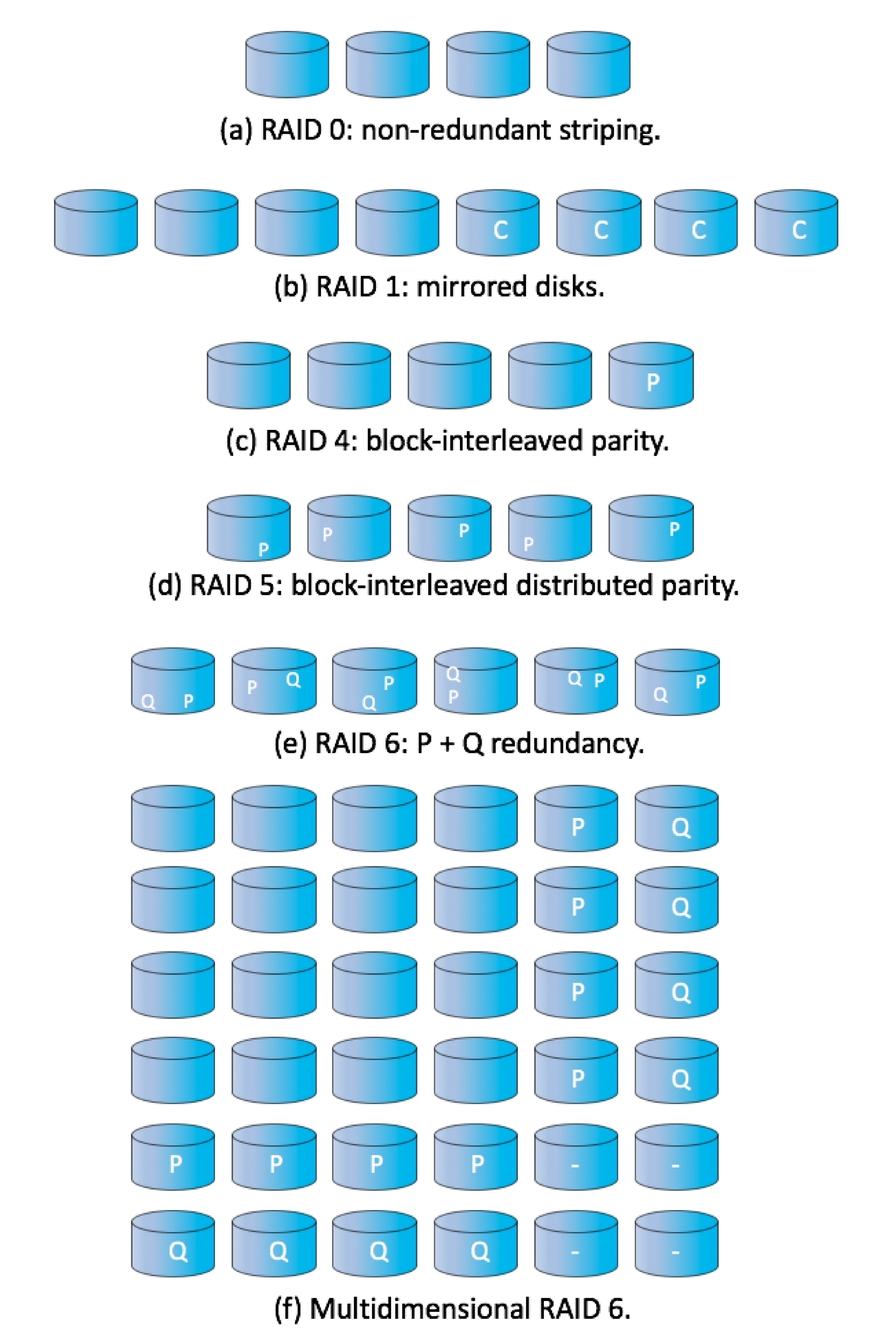
Server 1
Server 2
Server n
Switch B
Controller a
Disk array α
Controller b
Controller a
Disk array β
Controller a
Disk array γ
Controller b

# RAID – redundant array of inexpensive disks

- Disk striping uses a group of disks as one storage unit
- RAID is arranged into six different levels
- RAID schemes improve performance and improve the reliability of the storage system by storing redundant data
  - **Mirroring or shadowing (RAID 1) keeps duplicate of each disk**
  - Striped mirrors (RAID 1+0) or mirrored stripes (RAID 0+1) provides high performance and high reliability
  - **Block interleaved parity (RAID 4, 5, 6) uses much less redundancy**
- RAID within a storage array can still fail if the array fails, so automatic replication of the data between arrays is common
- Frequently, a small number of hot-spare disks are left unallocated, automatically replacing a failed disk and having data rebuilt onto them



# RAID Levels




(a) RAID 0: non-redundant striping.
(b) RAID 1: mirrored disks.
(c) RAID 4: block-interleaved parity.
(d) RAID 5: block-interleaved distributed parity.
(e) RAID 6: P + Q redundancy.
(f) Multidimensional RAID 6.

Hewlett Packard Enterprise

## BREAKDOWN OF COMMON RAID LEVELS

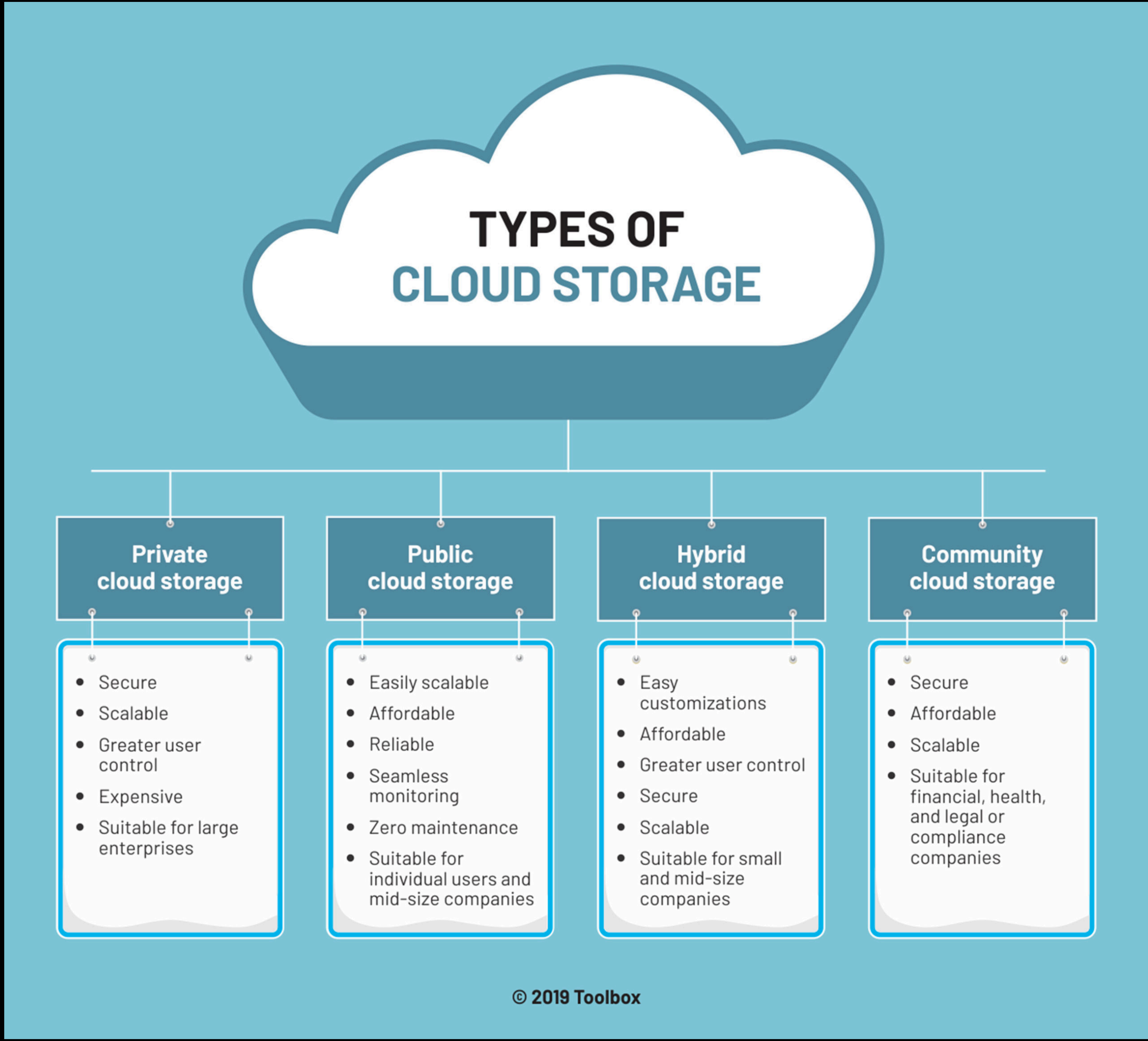
| RAID LEVEL | METHOD | HARDWARE / SOFTWARE | MINIMUM # OF DISKS | COMMON USAGE | PROS | CONS |
|---|---|---|---|---|---|---|
| JBOD | SPANNING | | 2 | INCREASE CAPACITY | COST-EFFECTIVE STORAGE | NO PERFORMANCE OR SECURITY BENEFITS |
| 0 | STRIPING | | 2 | HEAVY READ OPERATIONS | HIGH PERFORMANCE (SPEED) | DATA IS LOST IF ONE DISK FAILS |
| 1 | MIRRORING | | 2 | STANDARD APP SERVERS | FAULT TOLERANCE, HIGH READ PERFORMANCE | LAG FOR WRITE OPS, REDUCED STORAGE (BY 1/2) |
| 5 | STRIPING & PARITY | | 3 | NORMAL FILE STORAGE & APP SERVERS | SPEED + FAULT TOLERANCE | LAG FOR WRITE OPS, REDUCED STORAGE (BY 1/3) |
| 6 | STRIPING & DOUBLE PARITY | | 4 | LARGE FILE STORAGE & APP SERVERS | EXTRA LEVEL OF REDUNDANCY, HIGH READ PERFORMANCE | LOW WRITE PERFORMANCE, REDUCED STORAGE (BY 2/5) |
| 10 (1+0) | STRIPING & MIRRORING | | 4 | HIGHLY UTILIZED DATABASE SERVERS | WRITE PERFORMANCE + STRONG FAULT TOLERANCE | REDUCED STORAGE (1/2), LIMITED SCALABILITY |

### What Happened to 2-4 and 6-9?

The RAID levels described above are the most common levels used in enterprise scenarios. The levels in between are highly specialized and only make sense in very specific scenarios.

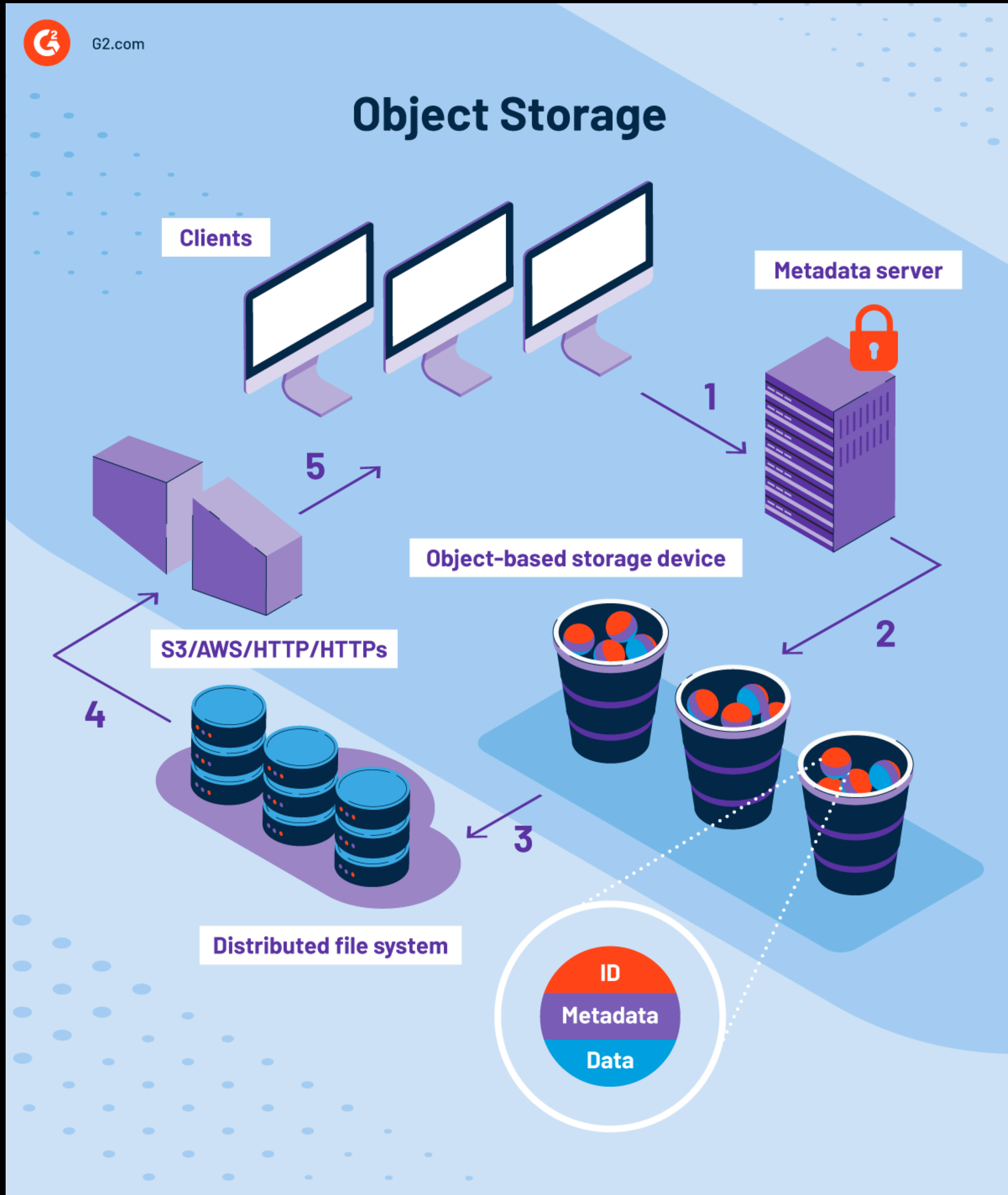
# Cloud Storage

- Similar to NAS, provides access to storage across a network
  - Unlike NAS, accessed over the Internet or a WAN to remote data center
- NAS presented as just another file system, while cloud storage is API based, with programs using the APIs to provide access
  - Examples include Dropbox, Amazon S3, Microsoft OneDrive, Apple iCloud
  - Use APIs because of latency and failure scenarios (NAS protocols wouldn't work well)



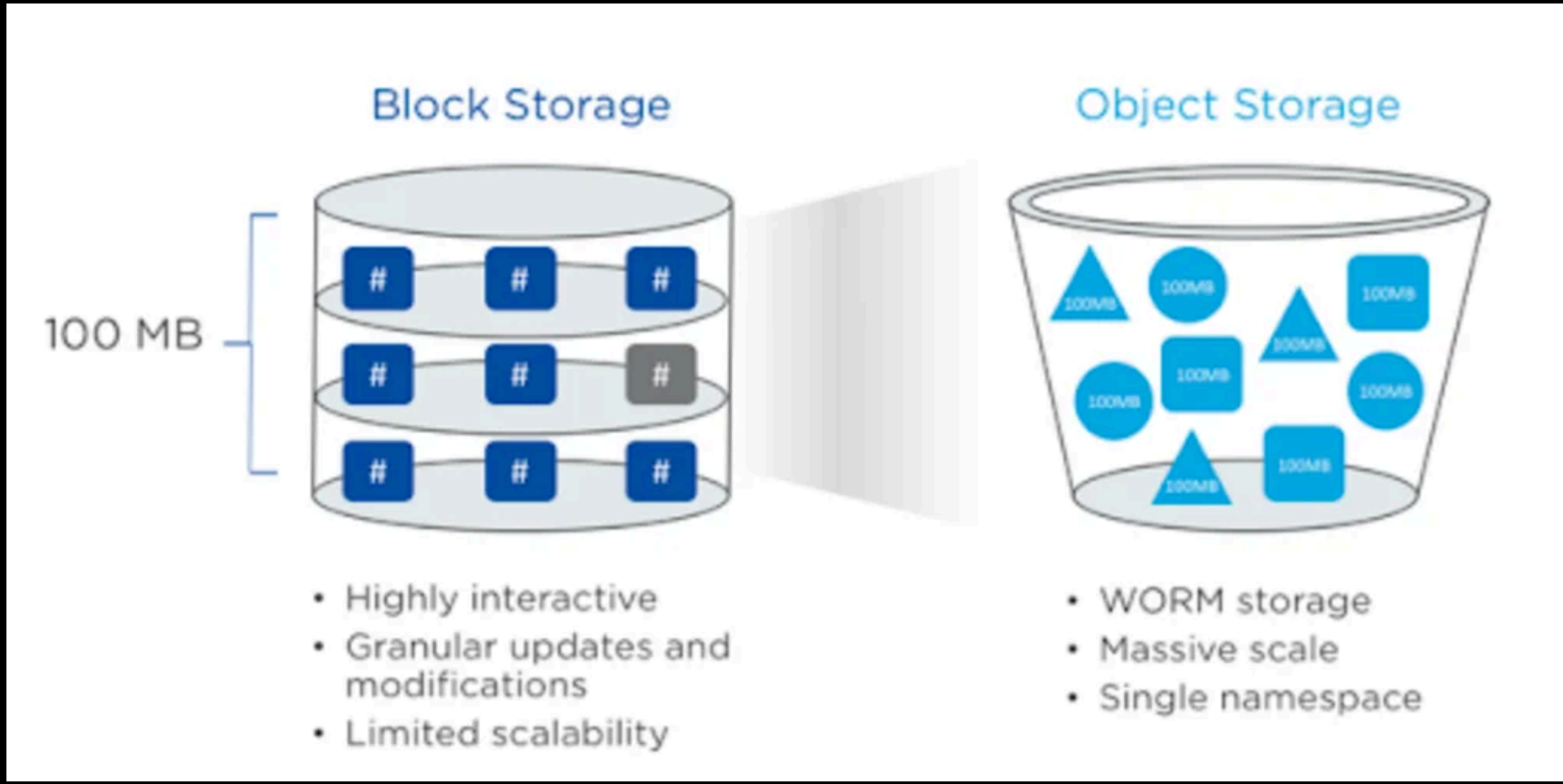
# Object Storage

- General-purpose computing, file systems not sufficient for very large scale
- Another approach – start with a storage pool and place objects in it
  - Object just a container of data
  - No way to navigate the pool to find objects (no directory structures, few services
  - Computer-oriented, not user-oriented
- Typical sequence
  - Create an object within the pool, receive an object ID
  - Access object via that ID
  - Delete object via that ID



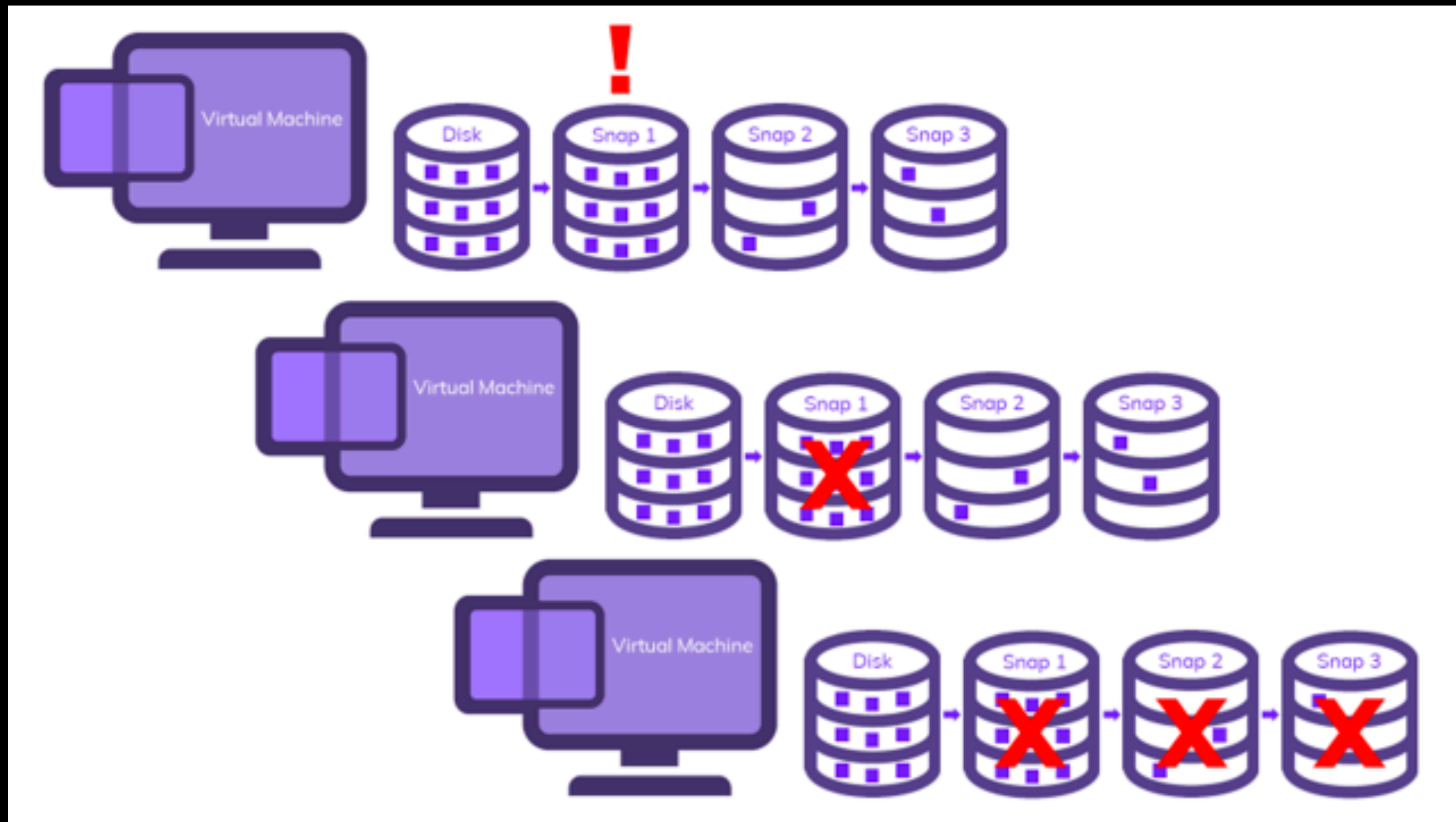
# Object Storage (Cont.)

- Object storage management software like Hadoop file system (HDFS) and Ceph determine where to store objects, manages protection
  - Typically by storing N copies, across N systems, in the object storage cluster
  - **Horizontally scalable**
  - **Content addressable, unstructured**



# Other Features on Storage devices

- Regardless of where RAID implemented, other useful features can be added
- **Snapshot**
- Replication is automatic duplication of writes between separate sites
  - For redundancy and disaster recovery
  - Can be synchronous or asynchronous
- Hot spare disk is unused, automatically used by RAID production if a disk fails to replace the failed disk and rebuild the RAID set if possible
  - Decreases mean time to repair



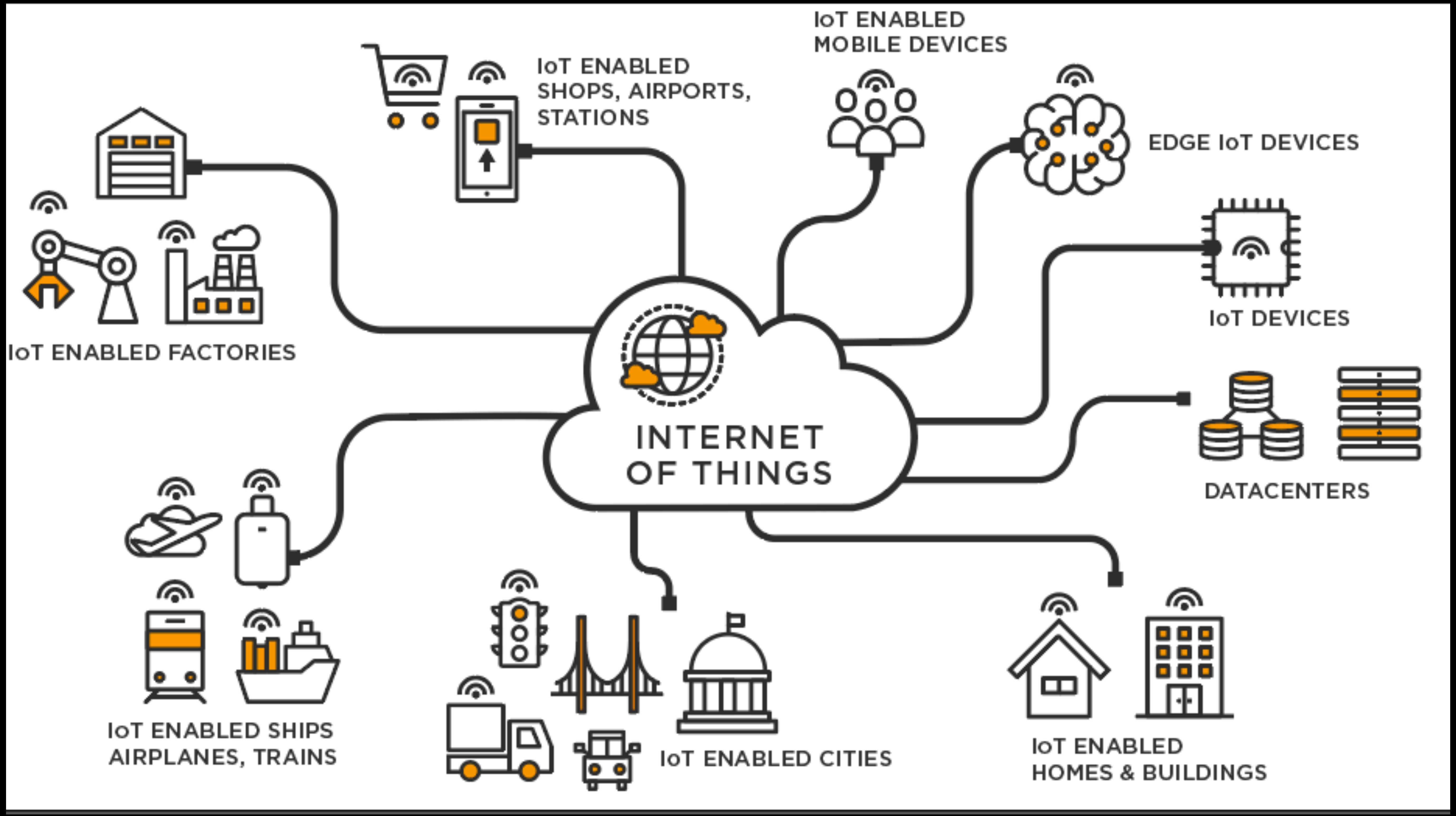

- Automatic failover
- Deduplication
- Tiering
- Virtual volume
- Provisioning

# Input / Output devices

- Display -> 8K

# Input / Output devices

- Display -> 8K
- IoT

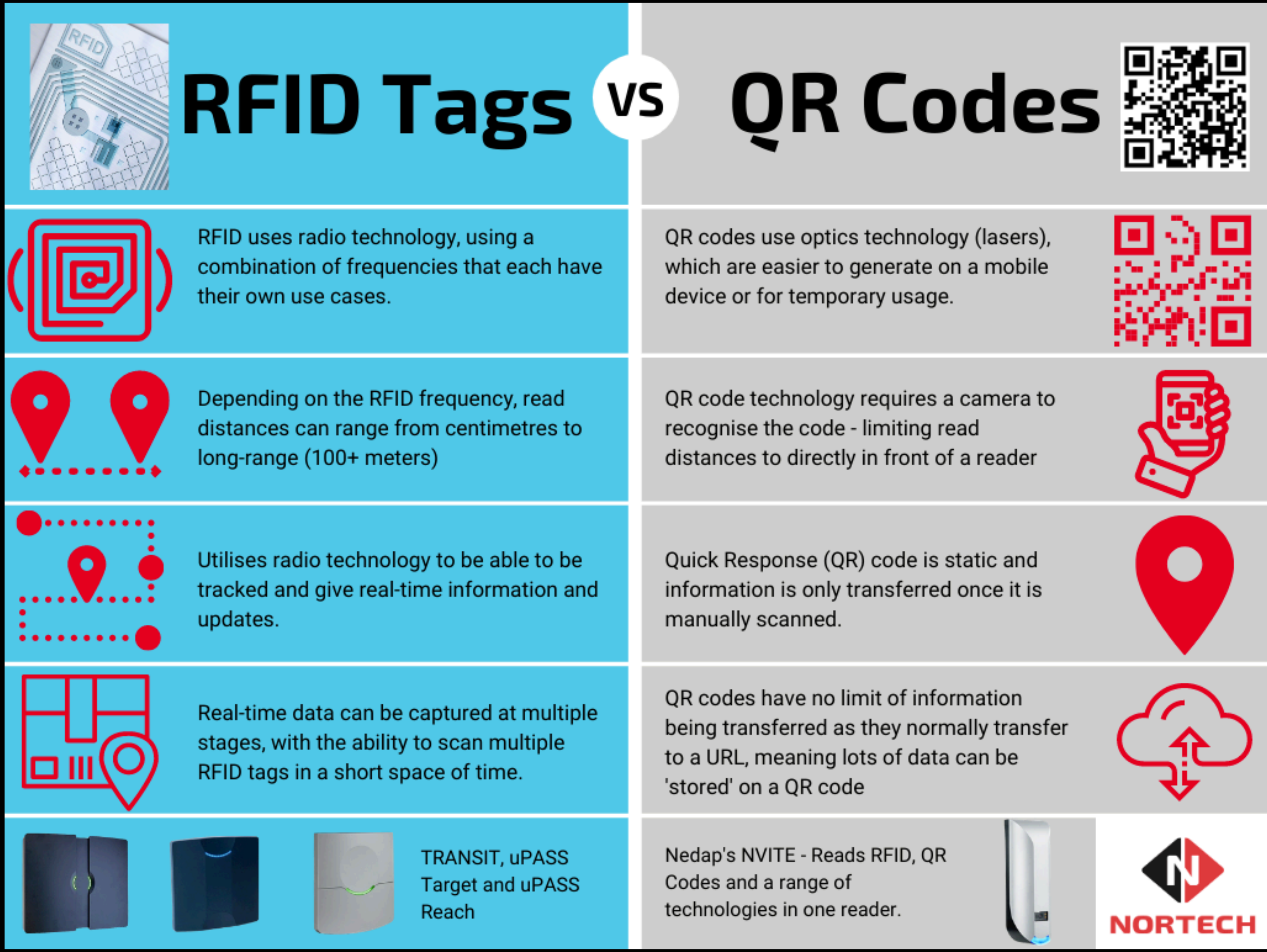


# Input / Output devices

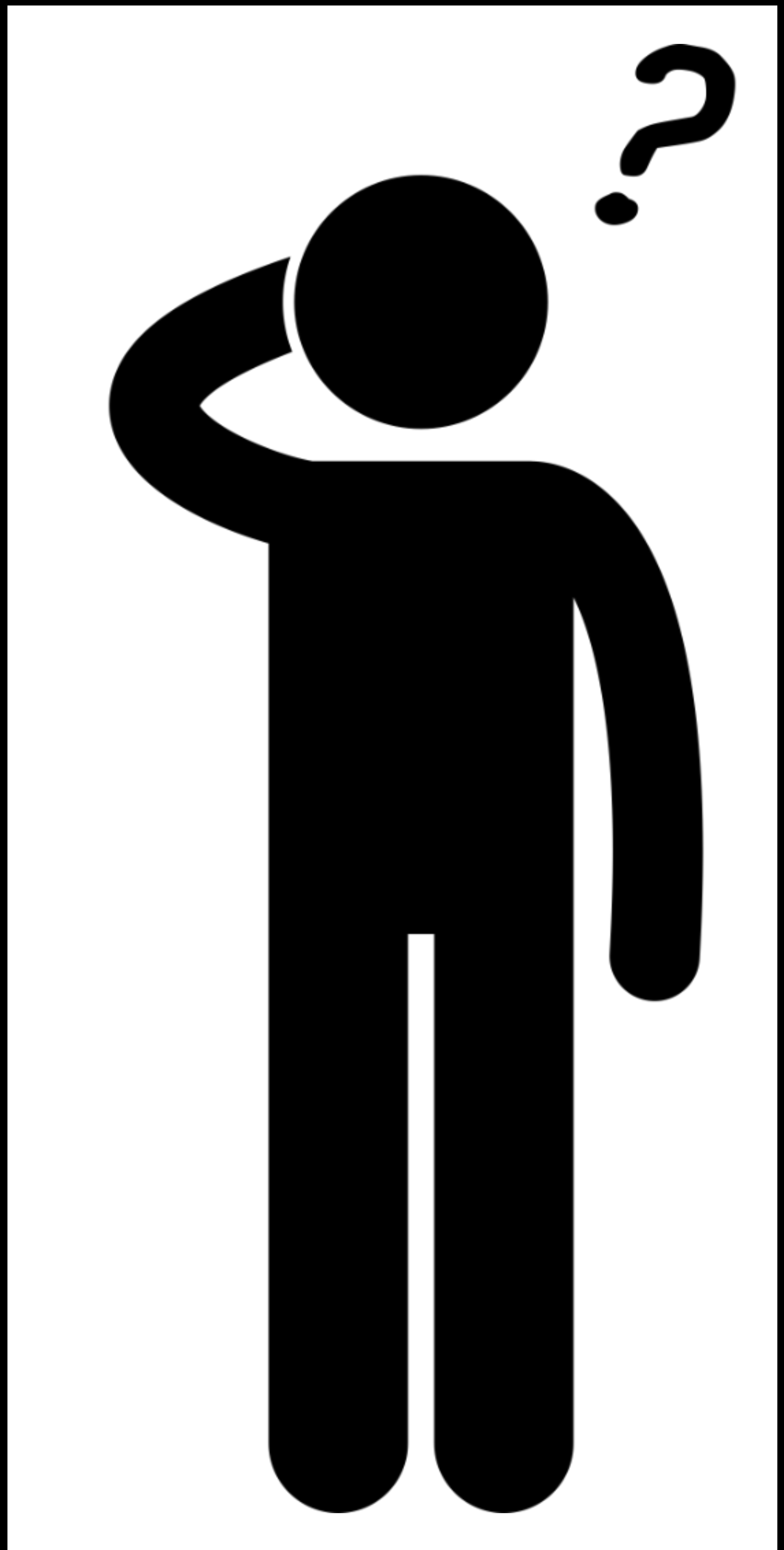
- Display -> 8K
- IoT
- Drones

# Input / Output devices

- Display -> 8K
- IoT
- Drones
- Smart thing
- RFID
- Tracking



## There are several key trends and examples that have leveraged the mentioned technologies to improve their supply chain



But has anyone really used all these fancy technologies for improving their supply chain?

# Amazon is using Predictive Analytics, 'Phygital' offerings and experimenting with emerging technologies to pioneer innovations in retail & technology spaces

CASE STUDY (1/4) : AMAZON

amazon.com


***Predictive Analytics***

- Based on a combination of previous orders, geography & other factors, Amazon will be able to predict orders even before they are shipped across
- This results in cost reduction, as well as the time given that it is able to predict orders of several people


***Phygital Offerings***

- Amazon Dash is a button that has bar code scanner and ability to place an order
- Through this, Amazon ensures that no customer ever forgets to order anything while at the same time the bar code scanner helps in asset tracking and ease of delivery for Amazon


***Emerging Technologies***

- Amazon Prime air is a delivery system designed to deliver goods within 30 mins
- Amazon has done several test runs globally and even filed a patent application in India
- The company is using AI for its web services arm (AWS) to analyze text and infer user sentiments

# DHL is using digital technologies to become a leader in logistics services

CASE STUDY (2/4) : DHL

***Augmented Reality***

- DHL is providing smart glasses to assist their employees in warehouse operations
- The functions embedded  include automatic bar-code scanning, color scanning etc.
- This ensures that workers are hands free improving their productivity
- DHL has reported a 25% increase in productivity on using these smart glasses

***Big Data Analytics***

- DHL is optimizing its last mile operations through Big Data Analytics
- It carries out Real time Route optimization of its vehicles
- ***DHL Smart Truck*** has:
  - Day to day optimization of initial tour plan based on shipment data
  - Dynamic routing system : changes routes depending on the current order and traffic situation
  - Cuts costs and improves $CO_2$ efficiency, for example by reducing mileage

# Adidas is using digitisation of supply chain to create a competitive advantage

CASE STUDY (3/4) : ADIDAS

***Product tracking using RFID***

- Adidas deployed RFID tracking across 400+ Russian stores to accurately inform customers about style and size availability. Resulted in ***increased NPS***[1]
- They have also used RFID to employ rigorous inventory management system

***Omni-Channel Sales***

- Adidas started "Click and Collect" programme which enabled user to book shoes online and collect them offline
- They started a reverse of this process with a programme called "Ship from store"

***Emerging Technologies***

- Consolidation of inventories between online and offline stores to rationalise their distribution
- Leveraging Data Analytics for portfolio optimisation – integrating customer information, RFID and beacons to track potential purchases & identify customers
- Adidas has partnered with Carbon Inc for 3-D Printing solutions

# In addition to IoT, Rio Tinto has been experimenting with other technologies like AI & VR

CASE STUDY (4/4) : ADIDAS

RioTinto

***IOT, analytics***

- RioTinto incorporates visualisation and collaboration tools to manage real time information operations Centre in Perth to manage all its mines, ports, rail roads
- They have also used using IoT to monitor tire burn out, and thereby manage tyre replacements proactively

***Virtual reality***

- As a  part of CRM initiative RioTinto has enabled virtual tour of its diamond mines in Diavik (Canada), in order to engage customers as well as other stakeholders in diamond industry better

***AI, IOT***

- Rio Tinto has implemented its first driverless train journey (which carry its Iron Ore), without any driver assistance
- They have enabled remote operations of blast hole drill systems, that enable an operator operate multiple drills from a single console remotely

# Increasing penetration of technological innovations across systems/ processes has high futuristic business impact

## TECHNOLOGY IN E-FULFILMENT: SUMMARY

### Autonomous conversations

1. Chatbots are going to replace 85 per cent of customer interactions by 2020, by end of 2018 faces and voices of customers will be recognized by digital assistants *- Gartner research predictions*

### Last Mile Autonomous Delivery technology

2. Truck platooning and autonomous trucks to be a reality by 2030; Semi-autonomous trucks will reach a penetration rate of 5 percent by 2030 *- Frost & Sullivan research*

3. Drone deliveries to open an $82 billion opportunity for the logistics and SCM industry over the next 10 years *- research paper presentation, e-commerce seminar, Singapore, Nov 2016*

### Artificial Intelligence

4. By 2020, 30% of all companies will employ AI to augment at least one of their primary sales processes *- Gartner research predictions*

# Group Discussion

Imagine, you've been asked by your manager to develop a roadmap for bringing digital change to your business.

Let's discuss **one way** a company can use digital in **each key business department** and explain why it'd be a benefit to introduce.

- Production
- Research and Development (often abbreviated to R&D)
- Purchasing
- Marketing (including the selling function)
- Human Resource Management
- Accounting and Finance
- Logistics and Supply Chain

# Digital Technologies are being employed as enablers to tackle supply chain related challenges

| *Digital enablers* | How Digital can overcome challenges ? |
| --- | --- |
| Big Data & Analytics | |
| Cloud Computing | |
| Internet of Things (IoT) | |
| Social Media | |
| Virtual & Augmented Reality | |
| Artificial Intelligence | |