CSC 411

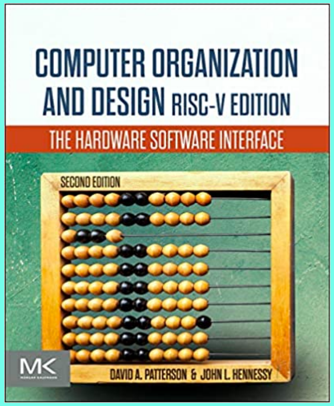
Computer Organization (Spring 2024)
Lecture 9: Computer systems

Prof. Marco Alvarez, University of Rhode Island

Disclaimer

The following slides are adapted from:

- Computer Organization and Design (Patterson and Hennessy)
  - The Hardware/Software Interface



The computer revolution

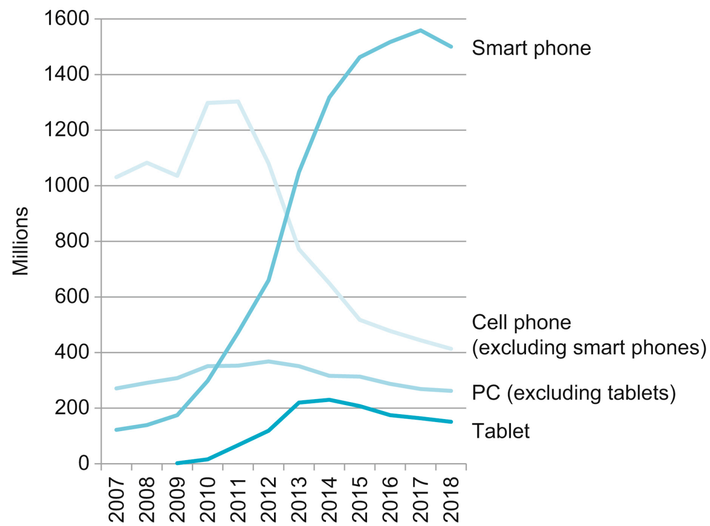
- Progress in computer technology
  - underpinned by domain-specific accelerators
- Novel applications (not long ago considered fiction):
  - generative AI, computers in automobiles, smartphones, human genome project, web, search engines
- Although common hardware technologies …
  - different design requirements

`Computers in all forms are now pervasive`

Classes of computers

- Personal computers (desktops/laptops)
  - low cost, general purpose, variety of software
  - subject to cost/performance tradeoff
- Server computers
  - network based, large workloads
  - high capacity, performance, reliability
  - range from small servers to building sized supercomputers (high-end scientific and engineering calculations)
- Embedded computers
  - largest class
  - hidden as components of systems (Internet of Things)
  - stringent power/performance/cost constraints

The post PC era (# of manufactured devices)



The post PC era

- Personal Mobile Device (PMD) — taking over PCs
  - battery operated
  - connects to the Internet
  - hundreds of dollars
  - smart phones, tablets, electronic glasses
- Cloud computing — taking over servers
  - Warehouse Scale Computers (WSC) — e.g. Amazon AWS
  - Software as a Service (SaaS)
    - portion of software run on a PMD and a portion run in the Cloud
  - Amazon and Google

Units of measure

SI prefixes

Factor	Name	Symbol	Factor	Name	Symbol
10^{1}	deca	da	10^{-1}	deci	d
10^{2}	hecto	h	10^{-2}	centi	c
10^{3}	kilo	k	10^{-3}	milli	m
10^{6}	mega	M	10^{-6}	micro	µ
10^{9}	giga	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y
10^{27}	ronna	R	10^{-27}	ronto	r
10^{30}	quetta	Q	10^{-30}	quecto	q

SI prefixes are a set of 24 prefixes used in the International System of Units (SI) to indicate multiples and submultiples of SI units. They are based on powers of 10, and each prefix has a unique symbol.

Binary prefixes

The SI prefixes refer strictly to powers of 10. They should not be used to indicate powers of 2 (for example, one kilobit represents 1000 bits and not 1024 bits). The names and symbols for prefixes to be used with powers of 2 are recommended as follows:

kibi	Ki	2^{10}
mebi	Mi	2^{20}
gibi	Gi	2^{30}
tebi	Ti	2^{40}
pebi	Pi	2^{50}
exbi	Ei	2^{60}
zebi	Zi	2^{70}
yobi	Yi	2^{80}

High performance computing

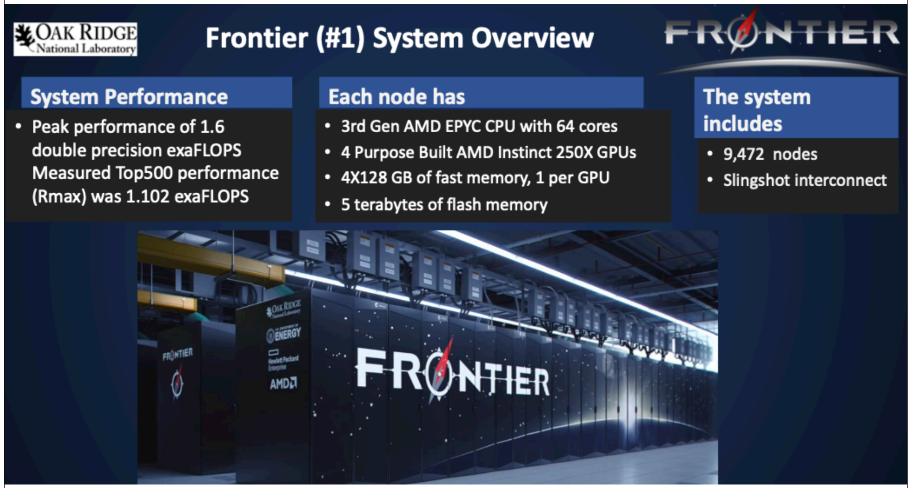
Top 500 project

- 500 most powerful computers in the world
- Updated twice a year
  - ISC'xy in June, Germany and SC'xy in November, U.S.

	TOP500 NOVEMBER 2023	Manufacturer	Computer	Country	Cores	Rmax [Pflops]	Power [MW]
1	Oak Ridge National Laboratory	HPE	Frontier HPE Cray EX235a, AMD EPYC 64C 2.0GHz, Instinct MI250X, Slingshot-11	USA	8,730,112	1,102	21.1
2	Argonne National Laboratory	HPE	Aurora* HPE Cray EX Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11	USA	4,742,808	585.3	24.6
3	Microsoft Azure	Microsoft	Eagle Microsoft NDv5 Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR	USA	1,123,200	561.2	
4	RIKEN Center for Computational Science	Fujitsu	Fugaku Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D	Japan	7,630,848	442.0	29.9
5	EuroHPC / CSC	HPE	LUMI HPE Cray EX235a, AMD EPYC 64C 2.0GHz, Instinct MI250X, Slingshot-11	Finland	2,069,760	309.1	6.0

credit: Berkeley's CS267 lectures

Top computer in the world

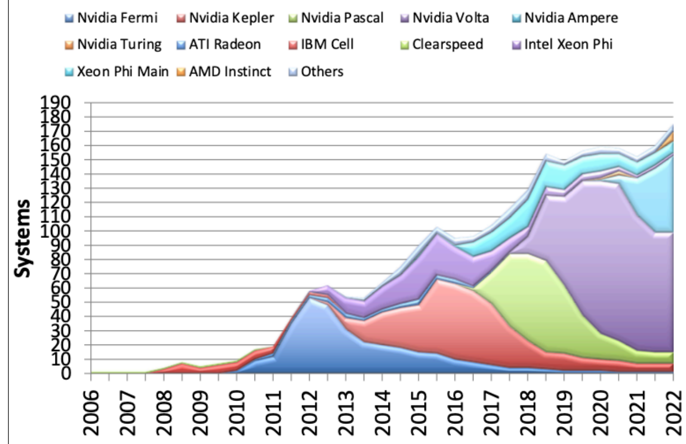


credit: Berkeley's CS267 lectures

Exascale applications at LBNL

credit: Berkeley's CS267 lectures

Heterogenous devices



2022: Over 1/3rd of all systems have **accelerators or co-processors**

The actual "performance share" is higher

credit: Berkeley's CS267 lectures

Abstractions

Seven great ideas

- Use **abstraction** to simplify design
- Make the **common case fast**
- Performance via **parallelism**
- Performance via **pipelining**
- Performance via **prediction**
- **Hierarchy** of memories
- **Dependability** via redundancy

Levels of program code

- High-level language
  - level of abstraction closer to problem domain
  - provides for productivity and portability
- Assembly language
  - textual representation of instructions
- Hardware representation
  - binary digits (bits)
  - encoded instructions and data

High-level language program (in C)

```
swap(int v[], int k)
{int temp;
   temp = v[k];
   v[k] = v[k+1];
   v[k+1] = temp;
}
```

Compiler

Assembly language program (for MIPS)

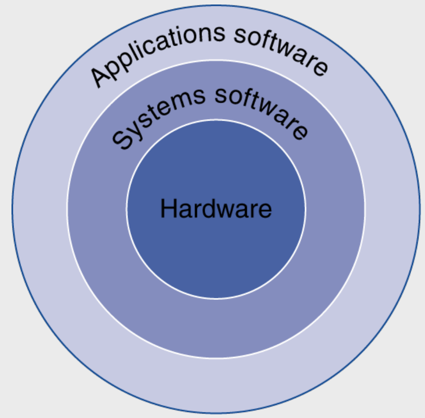
```
swap:
      muli $2, $5,4
      add  $2, $4,$2
      lw   $15, 0($2)
      lw   $16, 4($2)
      sw   $16, 0($2)
      sw   $15, 4($2)
      jr   $31
```

Assembler

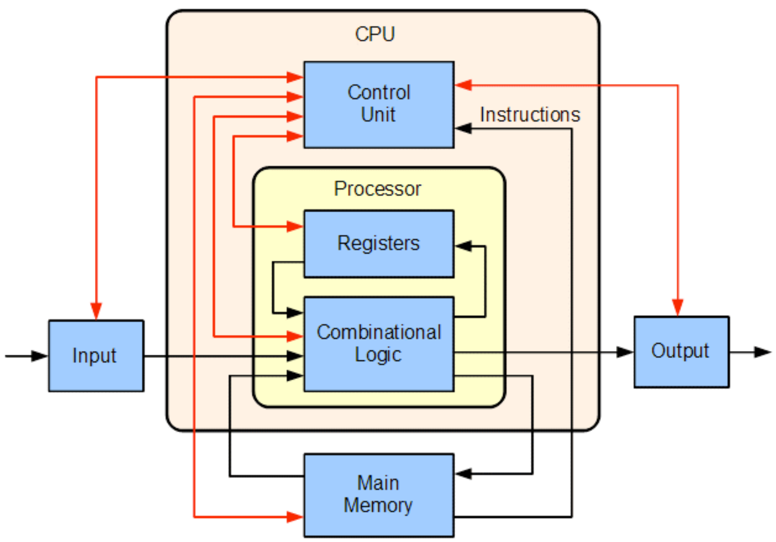
Binary machine language program (for MIPS)

```
00000000101000010000000000011000
00000000000110000001100000100001
10001100011000100000000000000000
10001100111100100000000000000100
10101100111100100000000000000000
10101100011000100000000000000100
00000011111000000000000000001000
```

Below your program



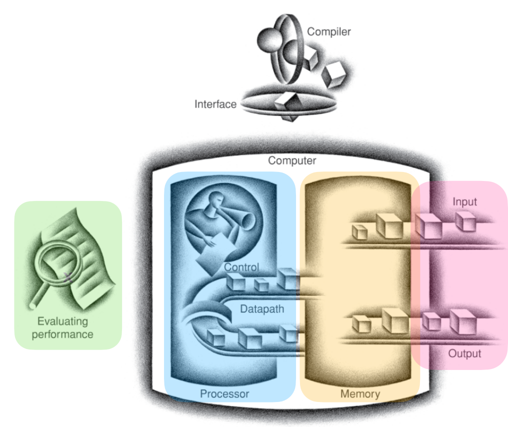
- Application software
  - written in high-level language
- System software
  - compiler: translates HLL code to machine code
  - operating system: service code
    - handling input/output
    - managing memory and storage
    - scheduling tasks & sharing resources
- Hardware
  - processor, memory, I/O controllers



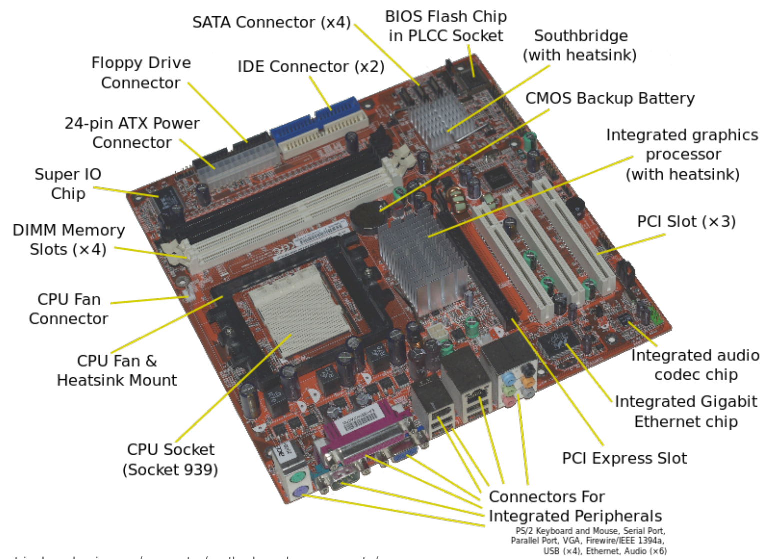
https://en.wikipedia.org/wiki/Computer_architecture

Components

- Same components for all kinds of computers
  - desktop, server, embedded
- Input/output includes
  - user-interface devices
    - display, keyboard, mouse
  - storage devices
    - hard disk, CD/DVD, flash
  - network adapters
    - for communicating with other computers

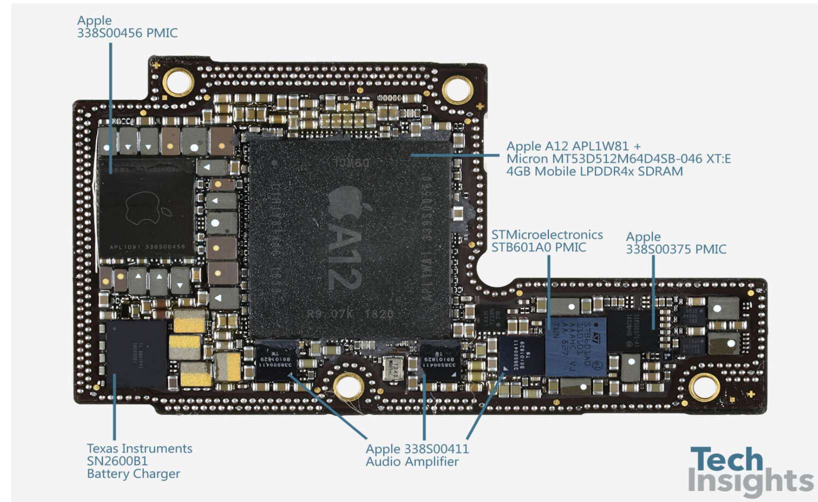


Opening the box



https://electricalacademia.com/computer/motherboard-components/

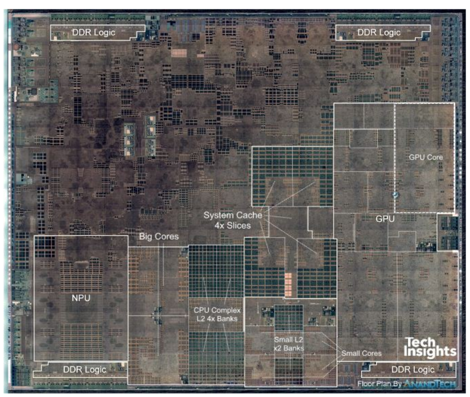
Opening the box



iPhone XS MAX (logic board)

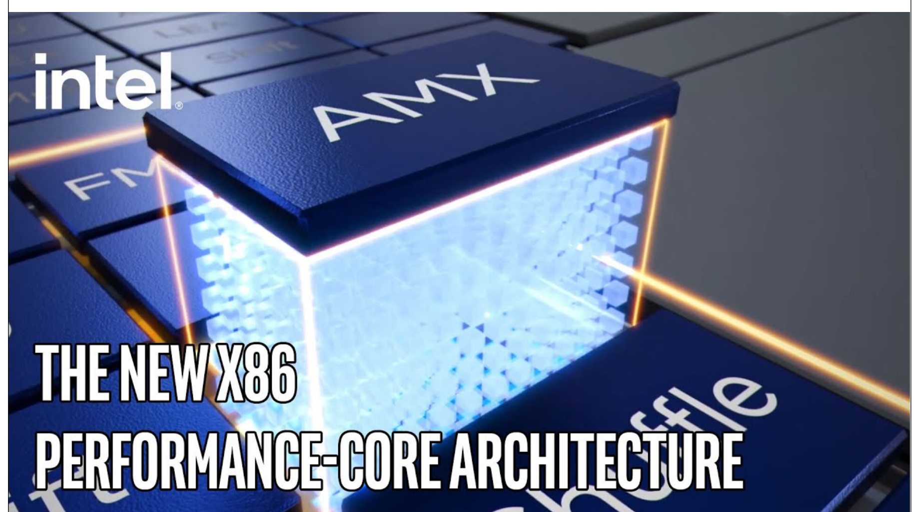
Looking inside

- Datapath
  - performs operations on data
- Control
  - sequences datapath, memory, ...
- Cache memory
  - small fast SRAM memory for immediate access to data



A12 processor

Intel's x86 core



https://www.youtube.com/watch?v=ijTRvIQV7bE

Chip Manufacturing

Technology trends

- Electronics technology continues to evolve
  - increased capacity and performance
  - reduced cost

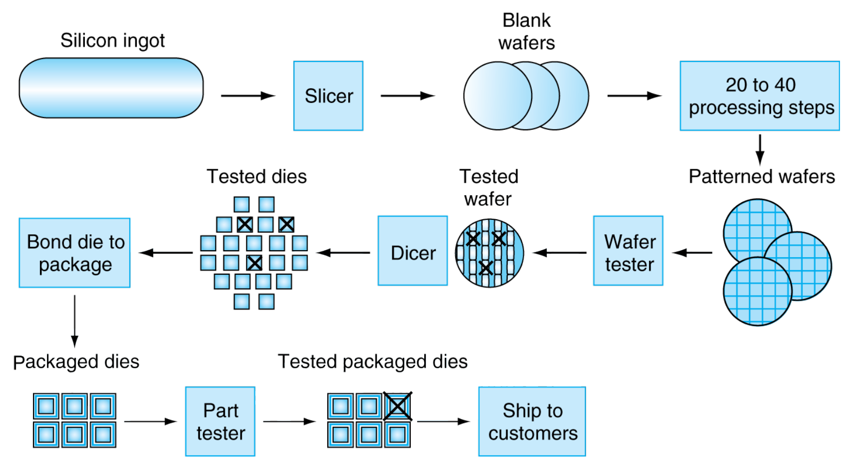
Year	Technology used in computers	Relative performance/unit cost
1951	Vacuum tube	1
1965	Transistor	35
1975	Integrated circuit	900
1995	Very large-scale integrated circuit	2,400,000
2020	Ultra large-scale integrated circuit	500,000,000,000

Semiconductor technology

- Silicon

- Add materials to transform properties
  - conductors of electricity
  - insulators
  - semiconductors (can conduct or insulate under specific conditions)

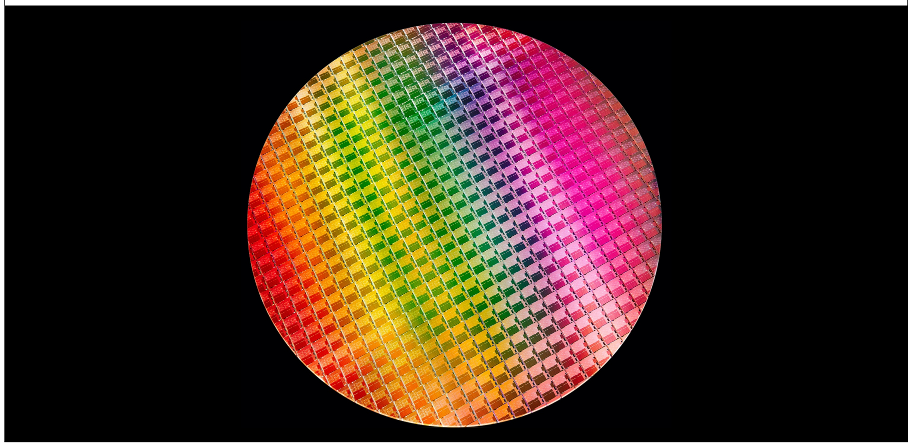
Manufacturing integrated circuits



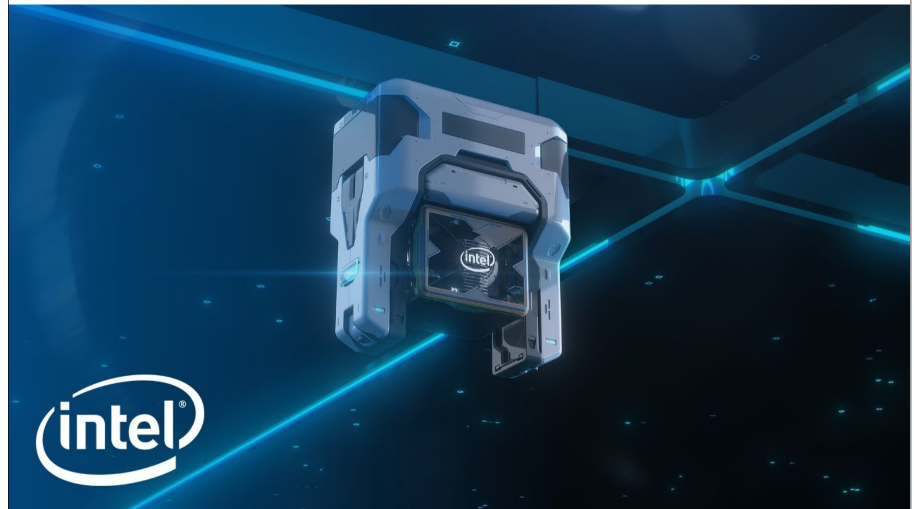
Yield: proportion of working dies per wafer

Intel® Core 10th Gen

- 300mm wafer (almost 1 foot), 506 chips, 10nm technology
- Each chip is 11.4 x 10.7 mm (a bit less than .5 in)



From sand to silicon



https://www.youtube.com/watch?v=_VMYPLXnd7E