# **Computer Architecture** Introduction

http://d3s.mff.cuni.cz/teaching/computer\_architecture/



#### Lubomír Bulej

bulej@d3s.mff.cuni.cz

**CHARLES UNIVERSITY IN PRAGUE** 

faculty of mathematics and physics

#### ... I'll be coding in Java, C#, Python, ..., JavaScript or PHP all day!

# Why do I need to know how a computer (or a processor) works ?



Computer Architecture, Introduction, summer 2022/2023

- Course credits ...
- It's mandatory ...









• I'll be coding web anyway...



#### ars TECHNICA

#### SUBSCRIBE 🔍 🗮 SIC

#### 

#### TECH-

# Steve Jobs: MobileMe "not up to Apple's standards"

Apple CEO Steve Jobs sent an internal e-mail to Apple employees this evening .

JACQUI CHENG - 8/5/2008, 5:19 AM

#### MailOnline

Home News U.S. | Sport | TV&Showbiz | Australia | Femail | Health | Science | Money | Video | Trave Latest Headlines | Royal Family | News | World News | Arts | Headlines | France | Pictures | Most read | Wires | Discounts

Australian Bureau of Statistics website CRASHES under the weight of traffic as millions attempt to complete their compulsory census form online - and avoid an \$1800 fine

- Australian Bureau of Statistics' census website has crashed online
- Millions have taken to social media to rant about the meltdown
- Residents have been met with error messages and unable to call hotline
- However, the online census is open until September 23

#### By CINDY TRAN FOR DAILY MAIL AUSTRALIA

PUBLISHED: 11:28 GMT, 9 August 2016 | UPDATED: 12:02 GMT, 10 August 2016

Millions of Australians attempting to complete the census online have been blocked from filling out their details after the website crashed - causing nationwide outrage.

The Australian Bureau of Statistics system descended into meltdown on Tuesday night, with millions of people taking to social media to rant their frustration.

Around 16 million people who are expected to log on to complete the compulsory survey on the night of August 9 have been met with error messages.

The online census has come to a crashing halt, leaving the website unavailable until further notice.

#### Firefox 3.0 released, servers overwhelmed

HOW-TO

DEALS

hs.

sh

By Peter Cohen

Macworld | JUN 17, 2008 12:24 PM PT

Mozilla has released Firefox 3.0, the final release of the newest version of its popular Web browser. The company may have been a bit unprepared for the



#### Government did not test health care site as needed

REVIEWS

Kelly Kennedy | USA TODAY Published 12:33 AM EDT Oct 25, 2013



The HealthCare.gov website has had problems with delays and dropped information. Lynne Sladky, AP

WASHINGTON-Not enough tests were performed on the HealthCare.gov website by the government and its contractors before the site was launched Oct. 1, a



- I'll be coding web anyway...
  - But it may be handy to know...
    - ... how things work in a computer, because it influences how operating systems, virtual machines, etc. work



- Course credits ...
- It's mandatory .
- I'll be coding web anyway...
  - But it may be handy to know...
    - ... how things work in a computer, because it influences how operating systems, virtual machines, etc. work
  - This will help me to ...
    - ... design and develop apps with more insight
    - ... diagnose and solve problems when (not if) they happen



#### • Cultivating mechanical sympathy...

... using a tool with an understanding how it operates best.

> "You don't have to be an engineer to be a racing driver, but you do have to have mechanical sympathy."

> > – Jackie Stewart, F1 racing driver

#### • Applied to computer science

- … improving program performance on modern CPUs
- ... better utilization of computing resources
- ... comparing performance of different computers and assess their suitability to a given task
- In a systematic fashion, not by trial and error



# Great ideas in computer architecture

- Design for Moore's law
- Use abstraction to simplify design
- Make the common cast fast
- Performance via parallelism
- Performance via pipelining
- Performance via prediction
- Hierarchy of memories
- Dependability via redundancy



# Technology



Computer Architecture, Introduction, summer 2022/2023

# **Processor and memory technology**

#### Transistor

- Basic building block
  - Discrete (a controllable switch) instead of analog (amplifier) application

#### Integrated circuit

- Multiple transistors on a single chip
  - Additional parts (capacitors, resistors, etc.)
- Better technology → smaller dimensions → higher level of integration → higher processor speed and higher memory capacity



# Processor and memory technology (2)



# **Processor and memory technology (3)**







# Processor

#### Key elements

- Data path (operates on data)
- Control (controls data path)
- Memory elements (registers and cache)

#### Intel Core i7-980X

- 6 cores, 12 MB L3 cache, clock frequency 3.33 GHz
- 32 nm technology, 248 mm<sup>2</sup>, 1.2 billion transistors



Source: intel.com

### Processor



#### **Cerebras Wafer-Scale Engine**

|                     | Gen1 WSE               | Gen2 WSE               |
|---------------------|------------------------|------------------------|
| Fabrication process | 16 nm                  | 7 nm                   |
| Silicon area        | 46,225 mm <sup>2</sup> | 46,225 mm <sup>2</sup> |
| Transistors         | 1.2 Trillion           | 2.6 Trillion           |
| AI-optimized cores  | 400,000                | 850,000                |
| Memory on-chip      | 18 GB                  | 40 GB                  |
| Memory bandwidth    | 9 PB/s                 | 20 PB/s                |
| Fabric bandwidth    | 100 Pb/s               | 220 Pb/s               |



Source: https://www.techpowerup.com/281313/cerebras-updates-wafer-scale-engine-on-7-nm-2-6-trillion-transistors-40-gb-onboard-sram-850-000-cores-12-wafer

# **Operating memory**

#### • Volatile

- Running programs and data
- Directly addressed by the processor
- Dynamic Random-Access Memory (DRAM)
  - Constant access time (tens of nanoseconds)
  - Bits stored as charge in capacitors
    - Needs periodic refresh (16 Hz typical)
  - Capacity in gigabytes

and a state of the state of the

# **Operating memory** (2)

#### Volatile

#### Static Random-Access Memory (SRAM)

- Implemented using two-state flip flops (requires 4 to 6 transistors per bit)
  - No need of periodic refresh
  - Significantly faster (units of nanoseconds), significantly lower density, significantly higher cost
- Processor caches and register
- Other kinds of processor-internal memory



18/42

#### Gordon Moore (\*1929)

- On of the founders of Intel
- Prediction: The number of transistors integrated on a single chip will double every 18 – 24 months
  - 1960s
  - Smaller transistors allow higher speeds and capacities
  - Often applied to other domains
    - Storage capacity, network bandwidth



# Growth of capacity per DRAM chip



Source: P&H



#### • Exponential growth in the last 40 years!

- Keeping Moore's "law" valid requires tremendous and continuous advances in technology
  - So far in a single domain (semiconductor transistors)
  - There are hard physical limits (quantum tunnel effect, waste heat, quantum noise)
- Compromises needed
  - Number of transistors does not correspond to computational power for sequential algorithms



# **Processor and memory technologies**

#### Impact of technology

- What computers will be able to do
- How fast will computers evolve
- Race to design a better computer
  - Embracing the latest in electronic technology

| Year | Technology   | Relative performance / unit cost |
|------|--|----------------------------------|
| 1951 | Vacuum tube  | 1                                |
| 1965 | Transistor   | 35                               |
| 1975 | Integrated circuit (low integration)                     | 900                              |
| 1995 | Integrated circuit (very large scale integration, VLSI)  | 2 400 000                        |
| 2013 | Integrated circuit (ultra large scale integration, ULSI) | 250 000 000 000                  |
|      |  |                                  |

# **Basic computer organization**



#### Computer

- input
- output
- memory
- processor
  - data path
  - control

#### Technology independent

Fits both today's and past computers



# Abstraction



Computer Architecture, Introduction, summer 2022/2023

# Abstraction

#### • Required to bridge semantic gaps

- From a concrete (technical) language to an abstract (general) language
- Expressing the same using more general terms while encapsulating internal details and preserving accuracy
  - More concise and compact expression
- "An abstraction is one thing that represents several real things equally well." (Edsger Dijkstra)



# Implementation

#### • The opposite of abstraction

- Concretization
- From computer architecture to concrete computer
- High-level language
  - Block diagrams, functional description of circuits
- Low-level language
  - Circuit diagrams connecting electronic components, masks for producing semiconductor elements in an integrated circuit
- "Machine code"
  - Physical realization of a computer



## From a user to an algorithm



#### • High-level programming language

```
void swap(unsigned int array[], unsigned int k) {
    unsigned int old = array[k];
    array[k] = array[k + 1];
    array[k + 1] = old;
}
```



# From an algorithm to a program





#### • Assembler representation for MIPS

```
swap:
    sll $a1, $a1, 2
    addu $a1, $a1, $a0
    lw $v0, 0($a1)
    lw $v1, 4($a1)
    sw $v1, 0($a1)
    sw $v0, 4($a1)
    jr $ra
```



#### Assembler representation for SuperH

```
swap:
    shll2 r5
    mov r4,r1
    add r5,r1
    mov.l @r1,r2
    add #4,r5
    add r5,r4
    mov.l @r4,r3
    mov.l r3,@r1
    rts
    mov.l r2,@r4
```



#### • Assembler representation for x86-64

```
swap:
    movslq %esi, %rsi
    leaq (%rdi, %rsi, 4), %rdx
    leaq 4(%rdi, %rsi, 4), %rax
    movl (%rdx), %ecx
    movl (%rax), %esi
    movl %esi, (%rdx)
    movl %ecx, (%rax)
    retq
```



# From a program to machine code





33/42

#### • Machine code for MIPS



#### • Machine code for SuperH



#### • Machine code for x86-64



# From power-on to running applications

#### • Firmware

- BIOS (Basic Input/Output System)
- Operating system loader
  - Boot sector
  - Boot loader
- Operating system
- User interface/desktop environment
- Application



# 100s of 1000s of lines of code

#### Application software

- Text editor, spread sheet, ...
- User interface libraries

#### System software

- Operating system
  - Input/output operations
  - Memory and storage management
  - Resource sharing
- Firmware

#### • Hardware

Processor, memory, I/O devices





# 100s of 1000s of lines of code



Source: https://informationisbeautiful.net/visualizations/million-lines-of-code (data as of 2016)





# **Abstraction layers in a computer**





# Beware: abstraction is (only) a tool!





42/42