Introduction to Multiuser Cluster Systems at LRZ
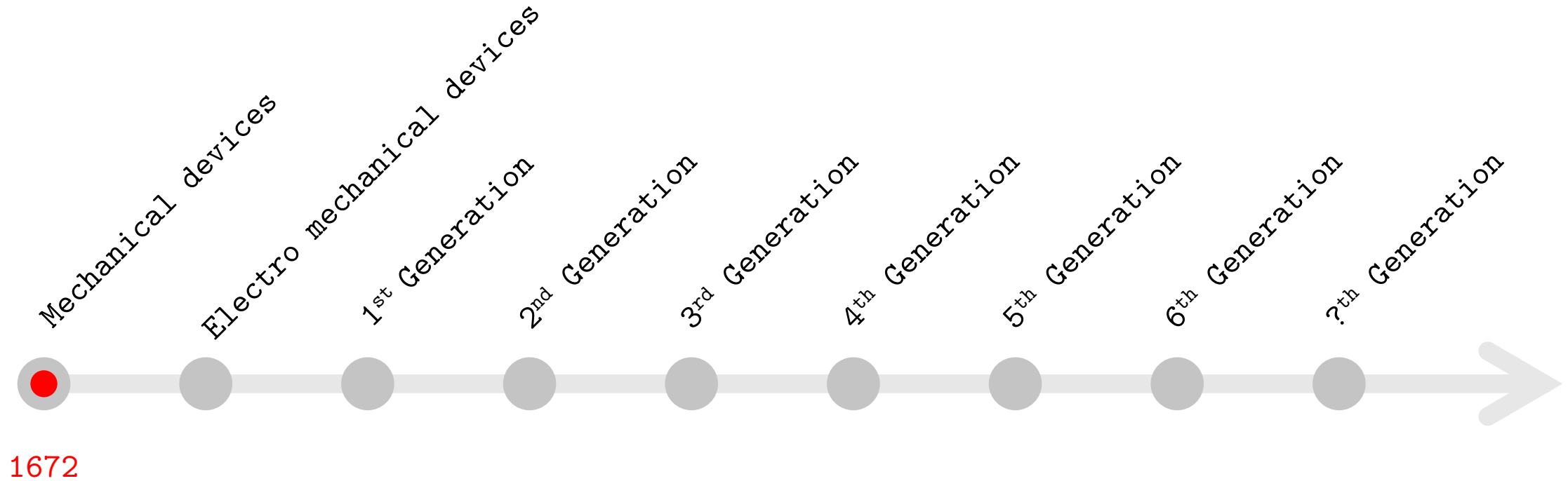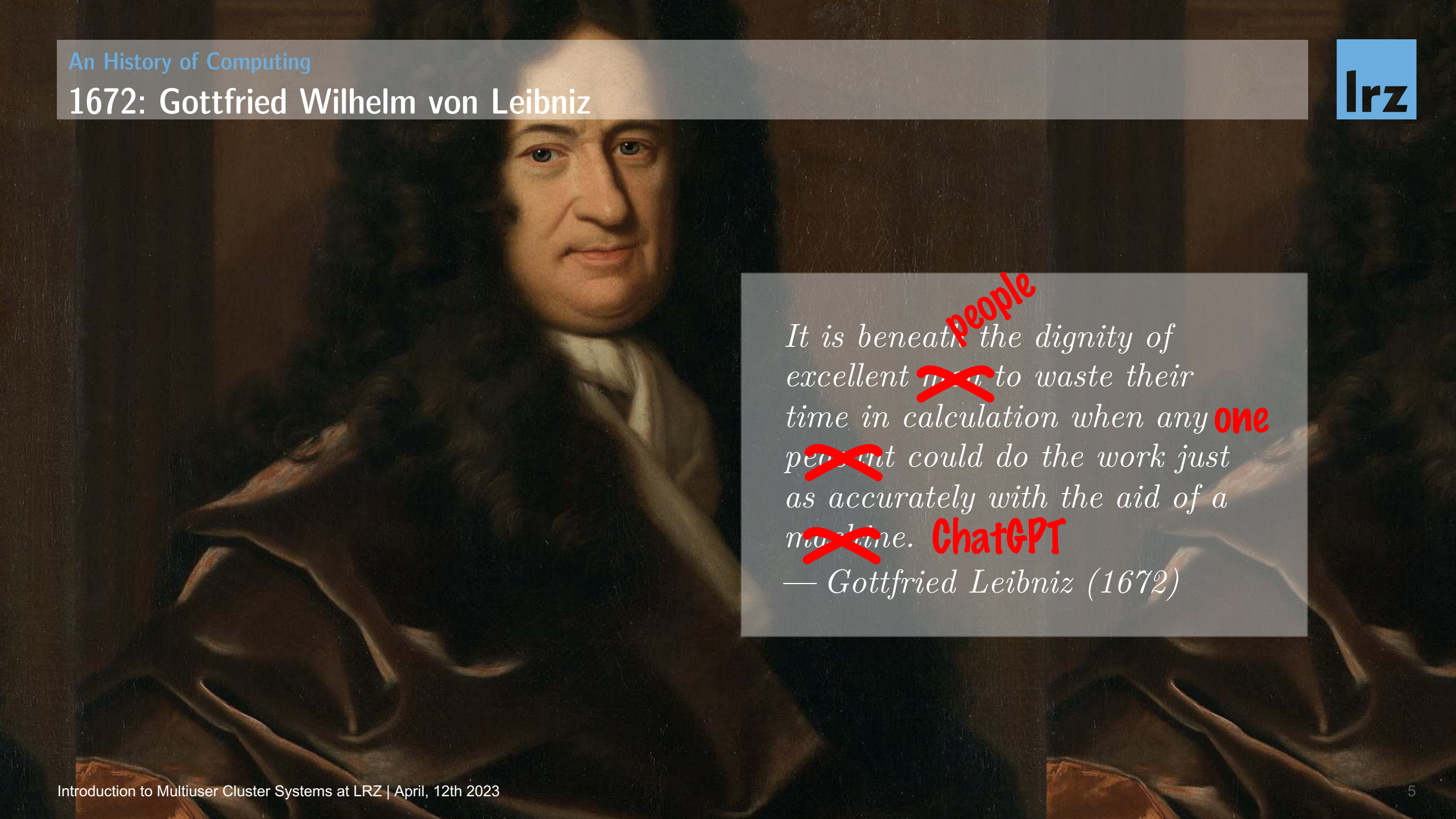
April, 12th 2023

- **Aim**: provide an introduction to multiuser cluster systems <u>in general</u> and to those operated at the <u>Leibniz Supercomputing Centre (LRZ)</u>, specifically

- You will probably benefit the most if you're not yet familiar with the LRZ HPC/HPDA/HPAI infrastructure, but plan to work with these systems in the future

- A majority of systems will be covered in more detail in dedicated sessions later this week

By the end of today's workshop, you should have a general understanding of multiuser HPC/HPDA/HPAI cluster systems and the basic skills to successfully interact remotely with such systems at LRZ

An History of Computing

The Generations of Computing Devices

Mechanical devices   Electro mechanical devices   1st Generation   2nd Generation   3rd Generation   4th Generation   5th Generation   6th Generation   ?th Generation
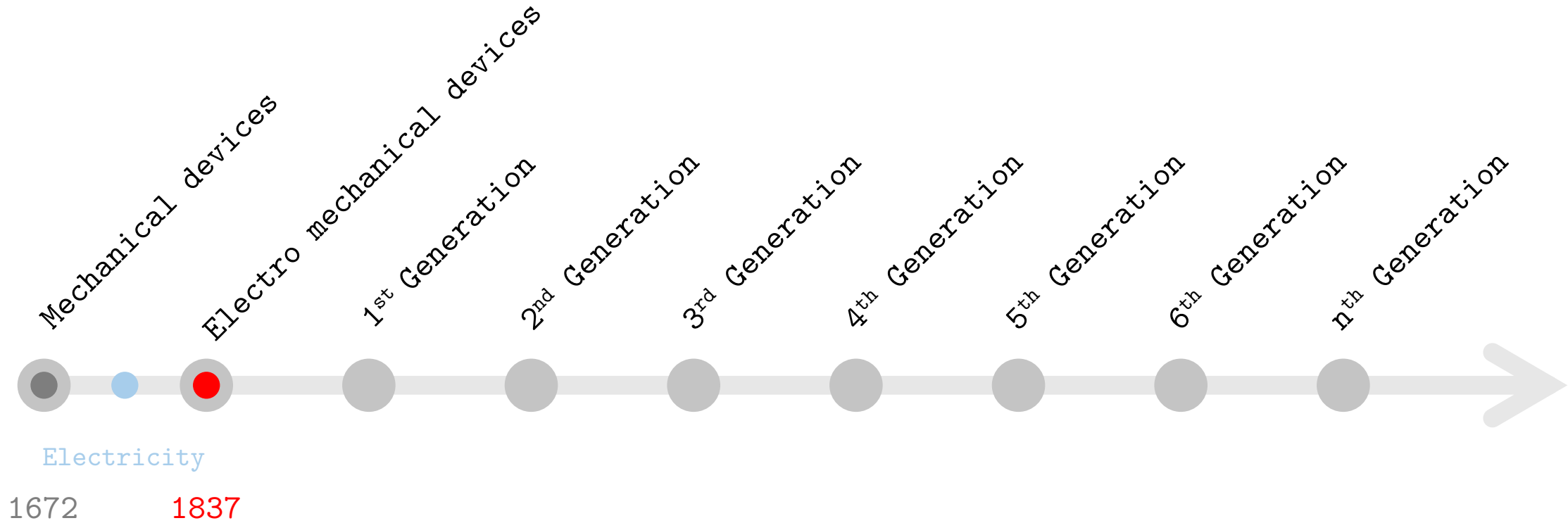
1672

lrz

It is beneath ~~the~~ people the dignity of excellent ~~men~~ to waste their time in calculation when any~~one~~ one ~~person~~ could do the work just as accurately with the aid of a ~~machine.~~ ChatGPT

— Gottfried Leibniz (1672)

1672 – 1673: Mechanical Device

The Generations of Computing Devices

Mechanical devices

Electro mechanical devices

1st Generation

2nd Generation

3rd Generation

4th Generation

5th Generation

6th Generation

n^{th} Generation

Electricity

1672

1837

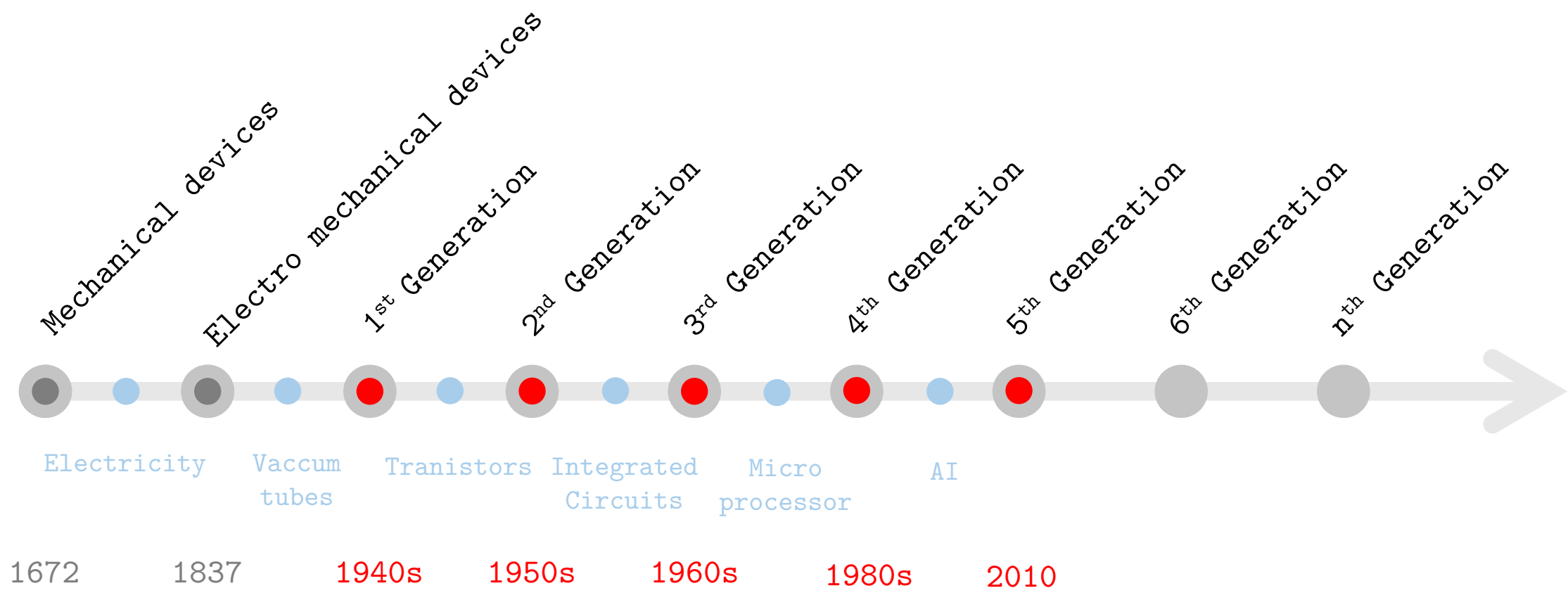# From the Jacquard loom to Electro Mechanical Devices

- **1752: Benjamin Franklin and his kite "discover" electricity** 🇺🇸
  - Prove that lightnings are a electrical discharge that can charge a conductor onto the ground
- **1804: The Jacquard Loom** 🇫🇷
  - Was a <u>mechanical</u> loom for cloth weaving
  - First demonstrated by Joseph Marie Jacquard in 1801.
  - Any number of the cards could be chained together into a continuous sequence, with each card corresponding to one row of the design
- **1837: Charles Babbage's Analytical Engine** 🇬🇧
  - First general-purpose computing device
  - <u>Electro mechanical</u> Device
  - Ada Lovelace as the first programmer 🇬🇧
- **Herman Hollerith: Punched card tabulating machine** 🇺🇸
  - Later became the "IBM Punchcards"
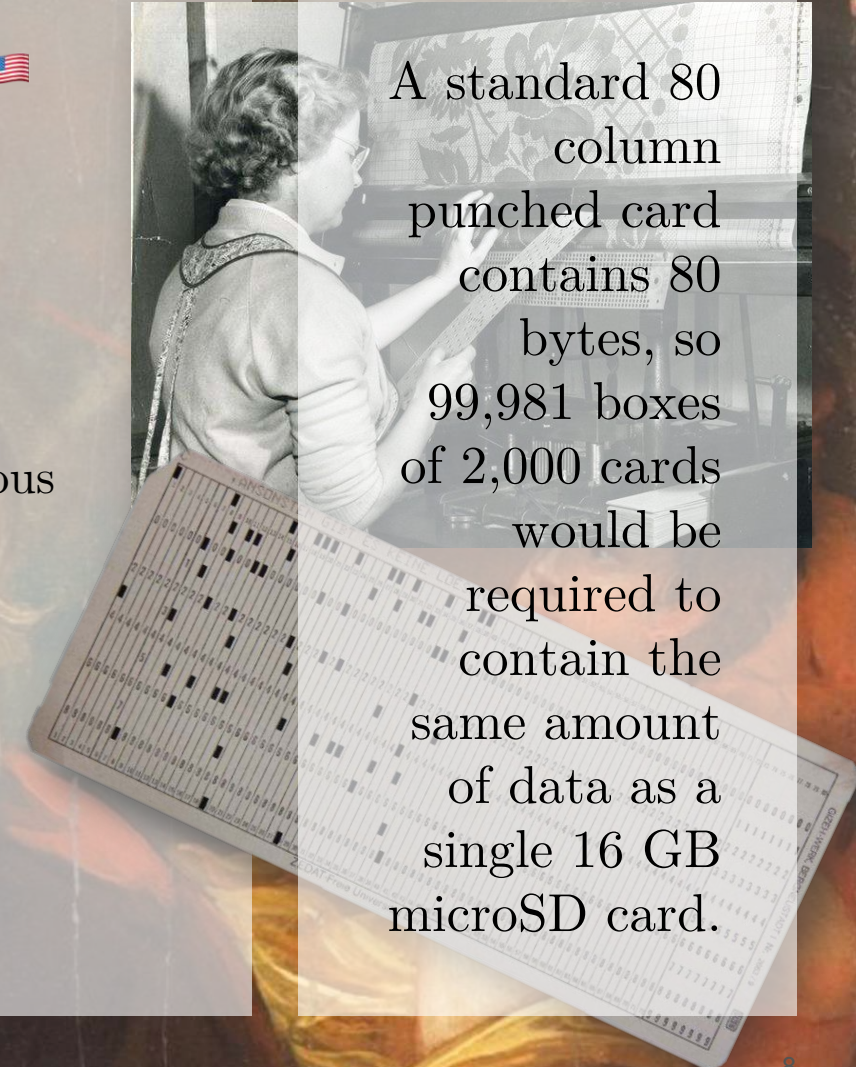
A standard 80 column punched card contains 80 bytes, so 99,981 boxes of 2,000 cards would be required to contain the same amount of data as a single 16 GB microSD card.

The Generations of Computing Devices

Mechanical devices

Electro mechanical devices

1st Generation

2nd Generation

3rd Generation

4th Generation

5th Generation

6th Generation

nth Generation

Electricity

Vaccum tubes

Tranistors

Integrated Circuits

Micro processor

AI

1672

1837

1940s

1950s

1960s

1980s

2010

1936: From Theory to the First Generation of Computers

- Alan Turing 🇬🇧

- On Computable Numbers, with an Application to the Entscheidungsproblem

- Paper that demonstrates the "Turing Machines" could perform solvable computation

- Describes the principles of the modern computer

- Also the foundation of Machine Learning

1940s-1950s: First Generation (Vaccum tubes)

- Used Vacuum tubes made of glass
  - Control flow of electricity between two electrodes: 0 || 1
  - Slow, unreliable, produced a lot of heat: Often would burn and would need to be replaced
- Heavy computers take up a full room
- Used for calculation, storage, and control purposes
- Main memory: Magnetic tapes and magnetic drums, IBM 650 would provide 4KB of storage
- No OS, no real programming language (machine code)
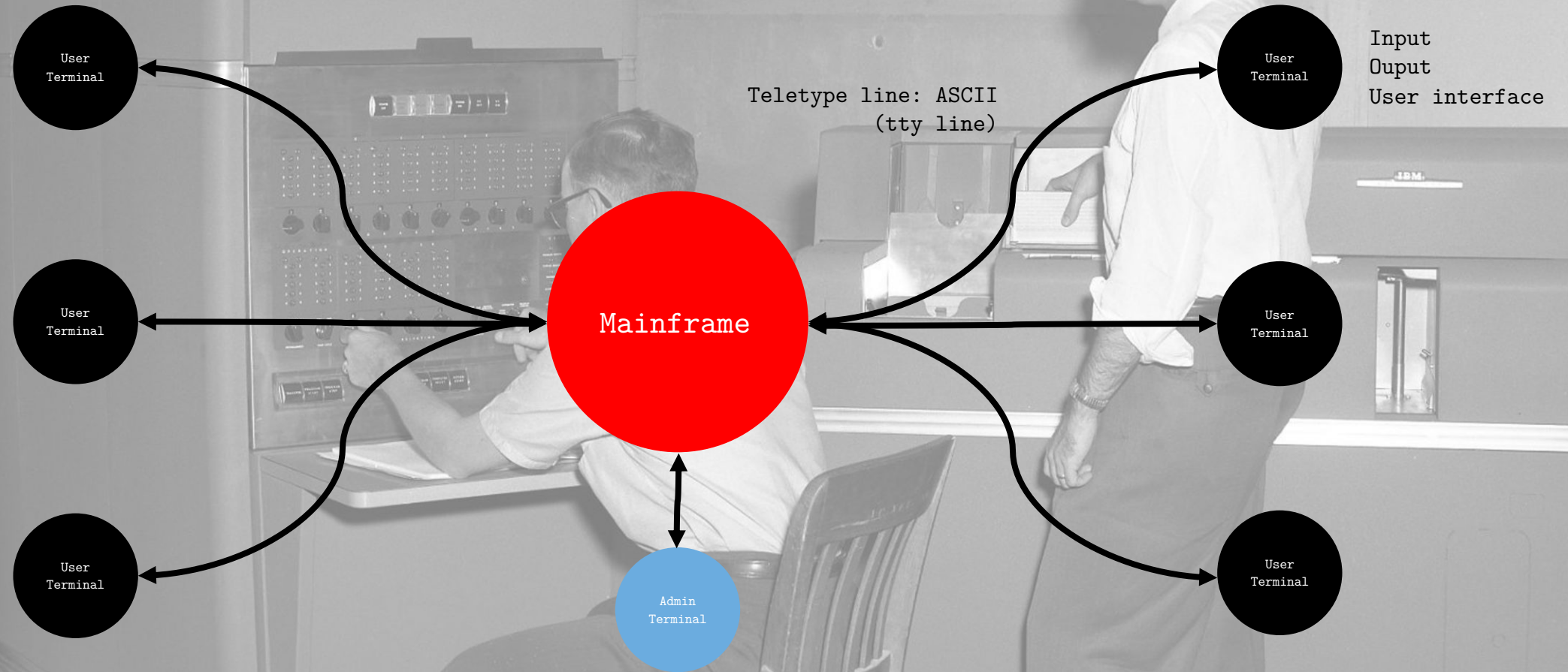- Example of machines: EDVAC, UNIVAC 1101, and IBM 650.
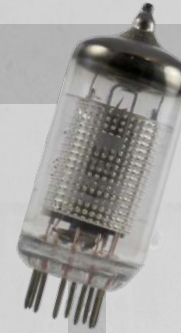
A mainframe is a large computer system that is usually used for multi-user applications. It is so expensive it needs to be share:

- time sharing
- space sharing
- batch processing
- ...

... all of which are still relevant.

Until the mid-to-late 1950s, the word "computer" referred to people who performed computations, not to machines.

1940s-1950s: First Generation, the centralized model

User
Terminal

User
Terminal

User
Terminal

Mainframe

Admin
Terminal

User
Terminal

User
Terminal

User
Terminal

Teletype line: ASCII
(tty line)

Input
Ouput
User interface

1950s-1960s: Second Generation (Tranistor)

- Use transistors instead of vacuum tubes
  - More reliable, smaller, and allow faster clock speeds
  - Transistors shaped the computer revolution and digital age: logical operations are performed by semi conductor devices
- Machine can store up to 2MB of data and run at 1 MHz
- Emergence of OS and multi-user support
- Basic networking capabilities
- Emergence of high-level programming language: FORTRAN (1956), ALGOL (1958), and COBOL (1959).
- Example of machine: IBM1400 series

The very first transistor — the foundational building block which almost all of modern civilization was built from — was created at AT&T's Bell Labs on December 23 1947.

A transistor is a semi conductor that uses an electric current to control the flow of electrons. It amplifies or switches electrical signals by enabling or preventing the flow of current. Multiple transistors combined form logic gates, arithmetic circuits, memories etc.

1960s-1980s: Third Generation (Integrated circuits)

- Transistors made smaller and packed into a silicon chip: towards integrated circuits
- Better speed and reliability
- Laguage: Becoming higher level: BASIC (Beginners All-purpose Symbolic Instruction Code).
- Example of machine: IBM System 360
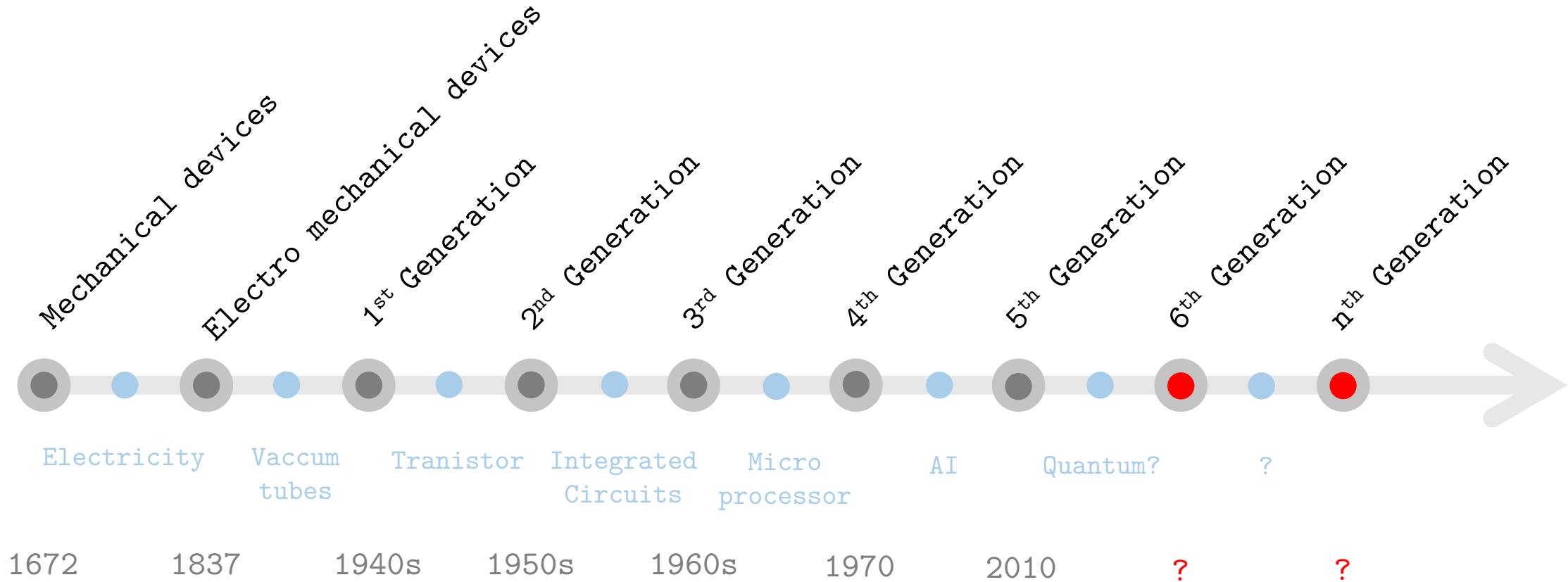- Beginning of minicomputers

1980s-2010: 4th Generation Computer (microprocessor)

lrz

- Macintosh 128k released in 1984
- Powered by a microprocessor (8 MHz) / 128 KB RAM
- 400 KB storage space on floppy
- First "real" Personal Computer (PC)
  - vs. IBM PC / Commodore 64
  - Comes with a screen, mouse, keyboard
  - Reaching a new audience: works without a manual
  - User friendly, cute, and adorable
- OS: System I (UNIX family, GUI)
  - Finder, Menu bar
  - Still the current HIG
- Application: MacPaint, MacWrite
- First affordable computer made for personal use ($2,500 ($6,500 in modern dollars))

- Key technologies include mobile devices (smartphones, tablets), cloud computing, social networks, high-speed wireless networks, IPv6 networking protocol, touchscreens, solid state storage, virtual/augmented reality, artificial intelligence

- Human like interaction and behaviour

  - Voice recognition

  - Computer vision

- Programming language: Very high level programming, Natural language

- Come in pocket size / wearables / Cloud only

- Digital twins, NVIDIA omniverse, Metaverse

The Generations of Computing Devices

Mechanical devices

Electro mechanical devices

1st Generation

2nd Generation

3rd Generation

4th Generation

5th Generation

6th Generation

nth Generation

Electricity

Vaccum tubes

Tranistor

Integrated Circuits

Micro processor

AI

Quantum?

?

1672

1837

1940s

1950s

1960s

1970

2010

?

?

Multiuser clusters...

... Haven't always been multiuser

... They are big and expensive: They need to be shared to cost-effectively serve a large number of concurrent users.

... They are meant to be used remotely

... System administrators take care of the system for the users

... can achieve much higher performance than individual systems by aggregating resources

... They provide fault tolerance through redundancy. If any part of the cluster fails, the rest of the cluster can continue operating

... Common architectures include...

   ... Server clusters: Multiple interconnected servers with shared storage and networking. Used for high availability and scalability.

   ... High-performance computing clusters: Powerful servers with fast interconnects, used for running highly parallel workloads.

   ... Cloud computing clusters: Massive clusters that run cloud platforms and services, accessed by many users over wide-area networks.

   ... Edge clusters: Clusters deployed at the edge (near users/devices) to support localized computing, storage and networking needs.
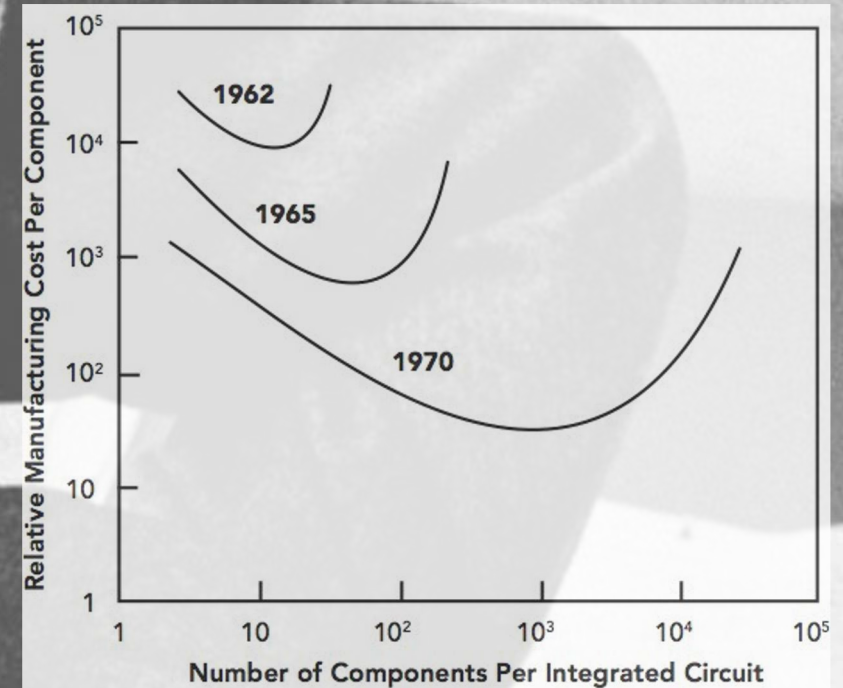
Moore's law

The number of transistors on integrated circuits is doubling about every ~~18 months~~ 2 years.
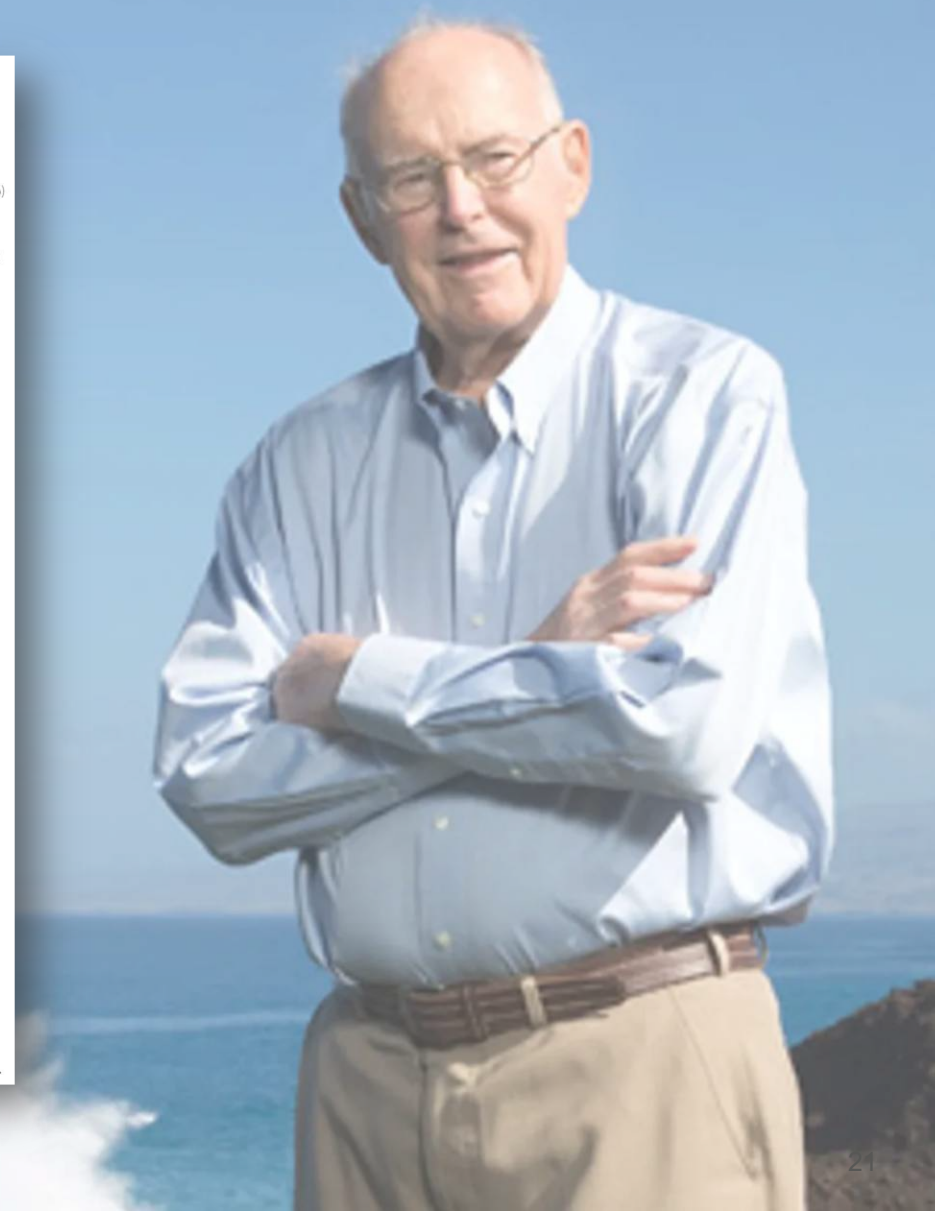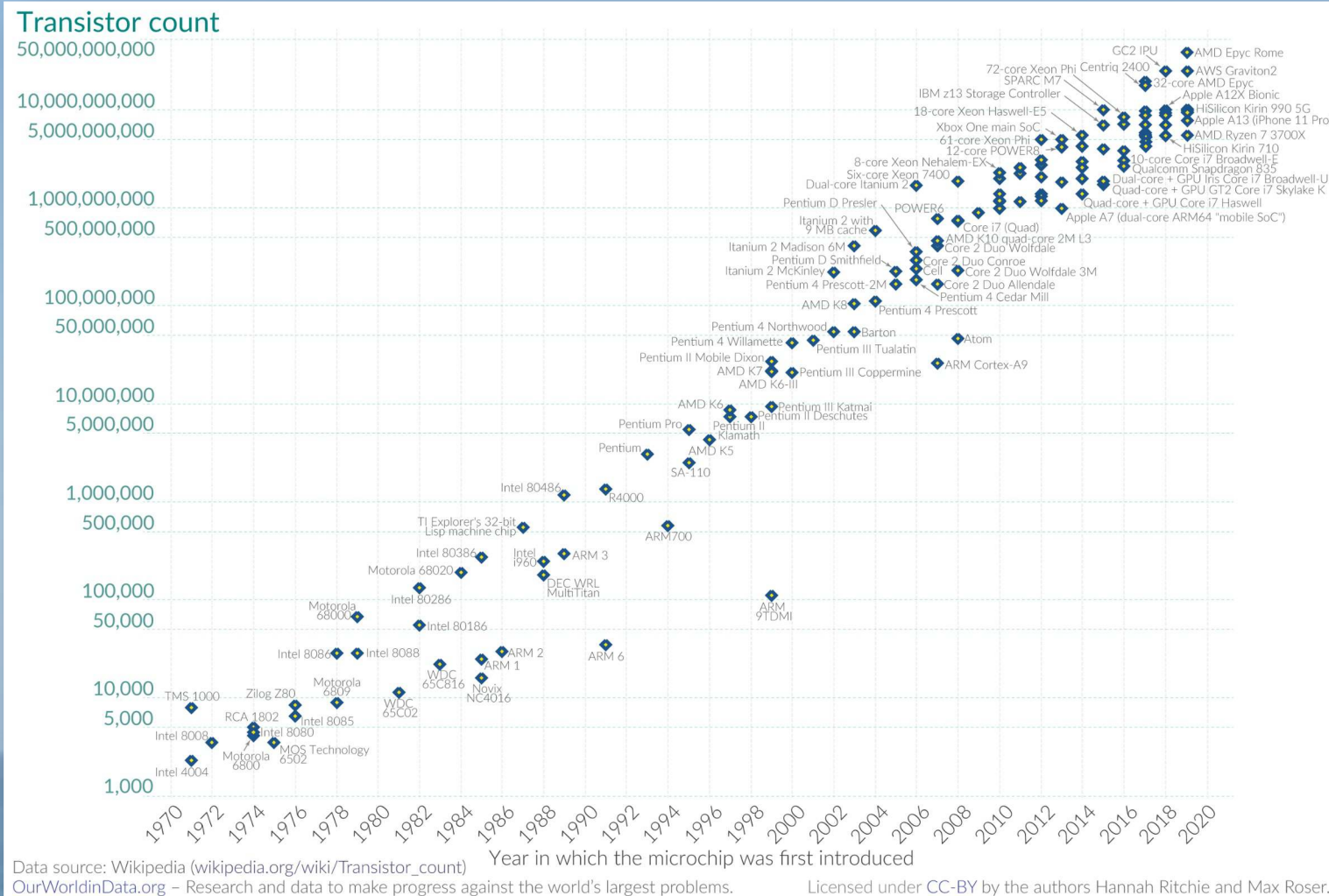
— Gordon Moore, 1965

Gordon Moore's curves on this plot show that development of the chemical printing technology makes more complex microchips the cheapest form of electronics.

Source: Gordon Moore.

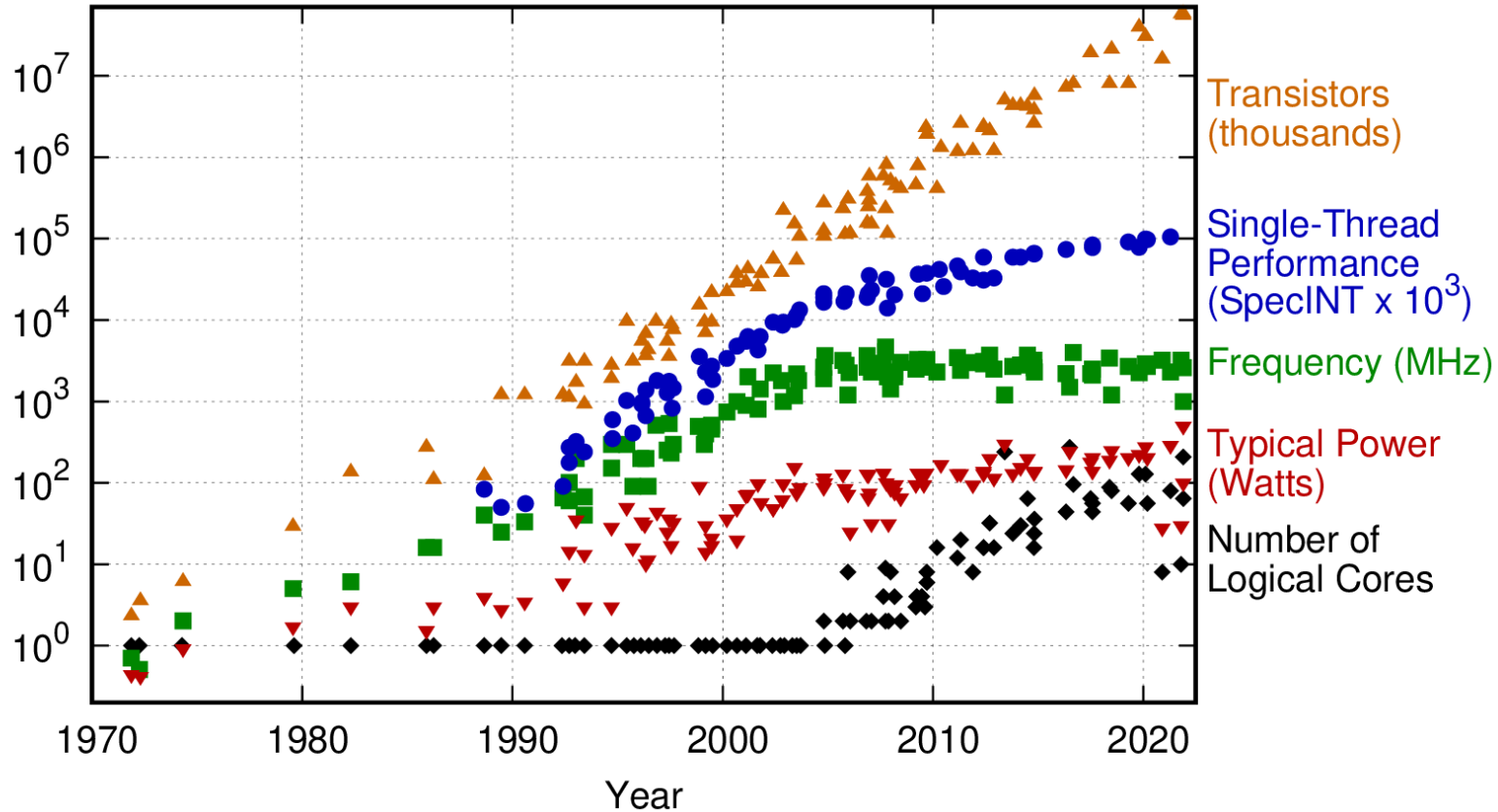Transistor count vs. Year in which the microchip was first introduced

Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)
OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

50 Years of Microprocessor Trend Data

Transistors (thousands)

Single-Thread Performance (SpecINT x 10^3)

Frequency (MHz)

Typical Power (Watts)

Number of Logical Cores

Year

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2021 by K. Rupp

- Mid 2000s: "heat death"

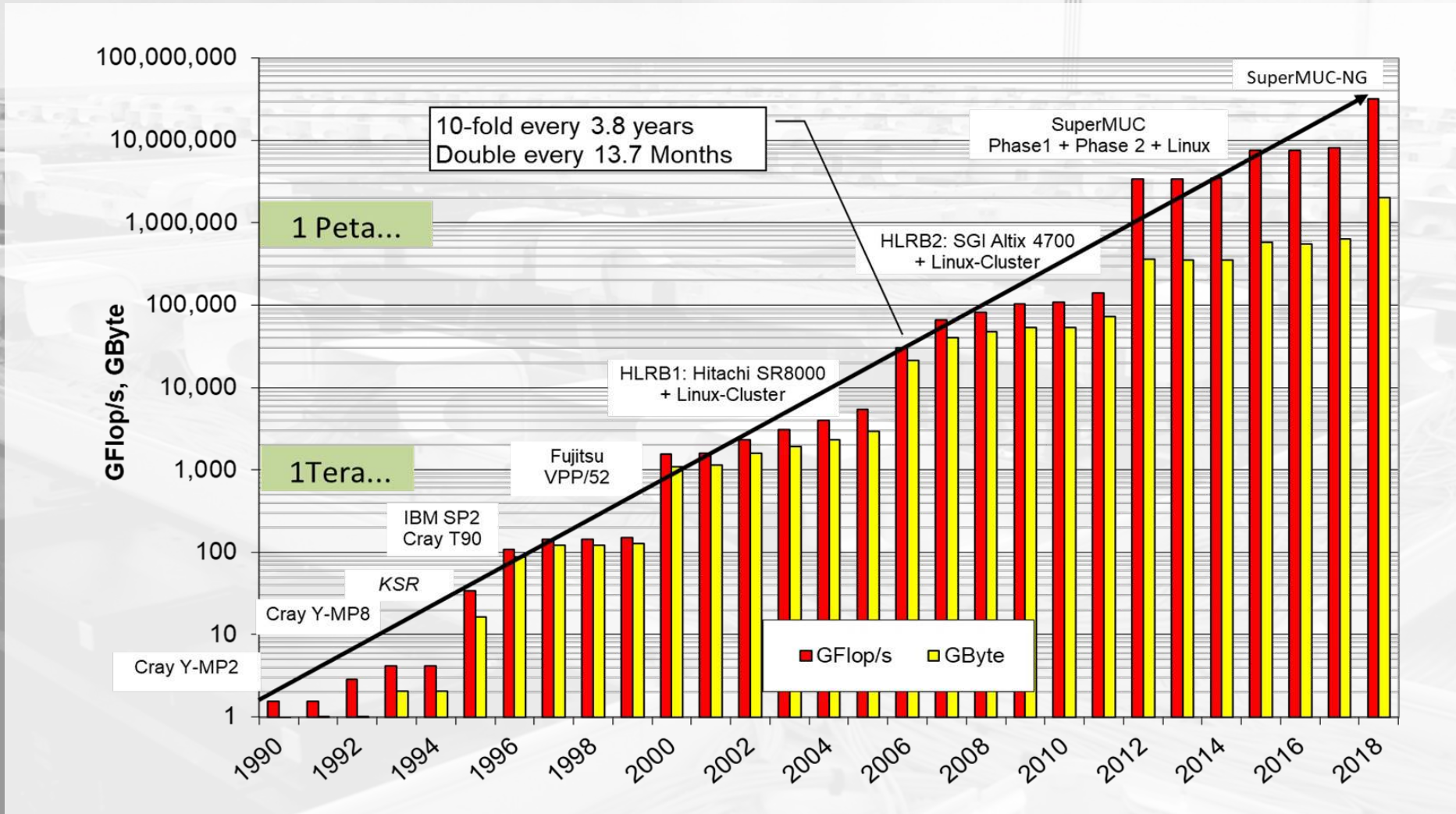- No more faster processors, only more of them.

- But: 2x3 GHz != 6 GHz

Is Moore's Law dead?
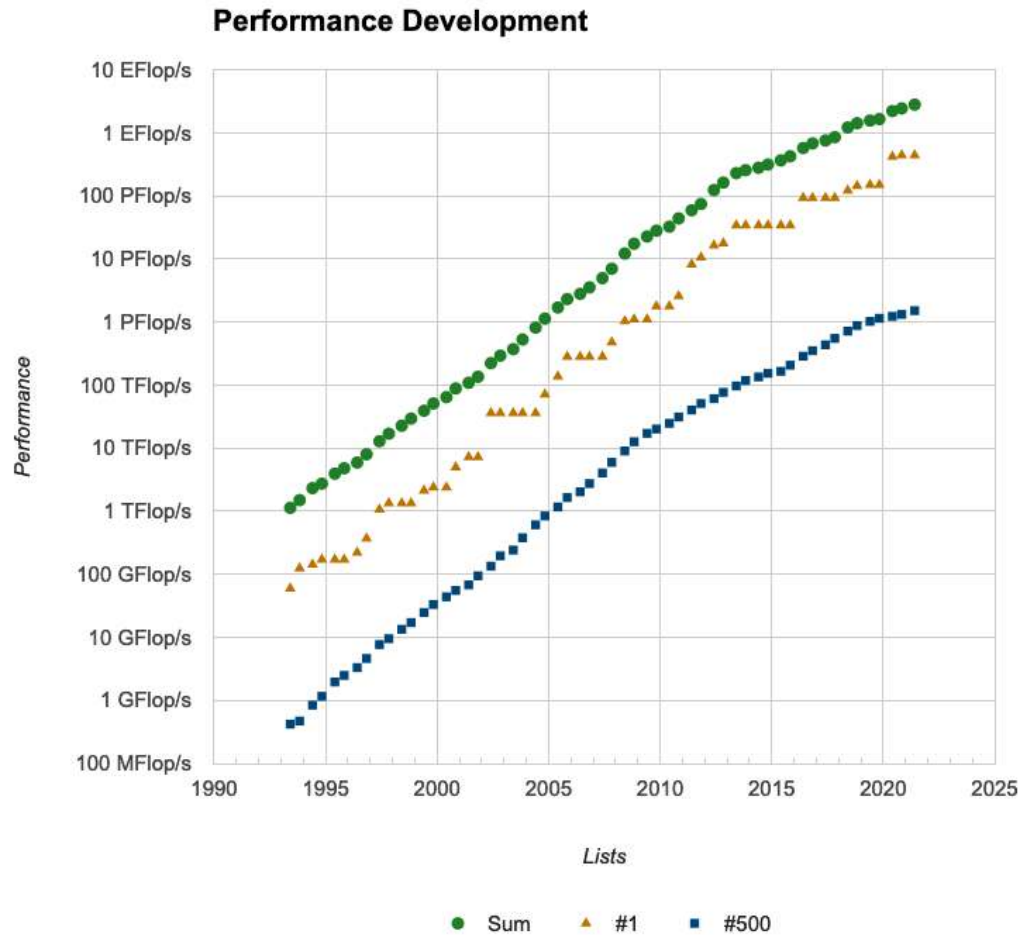
- **From #1 to #500:** 6-8 years

- **From #500 to Notebook:** 8-10 years

Evolution of Peak Performance and Memory

What is a Supercomputer or High-Performance Cluster... (Not)?

It runs Microsoft Windows?	😌 No, no worries
It will run my Excel spreadsheet?	😏 No!
It has overclocked high-speed processors?	🤨 No
The CPU runs faster than a desktop PC?	🤔 Not even
It has a large internal memory (RAM)?	😱 Usually not (except exceptions)
It will run my old tried and tested executable?	😔 Probably not
It will run my software without changes?	😵‍💫 Probably not
It will run my program with millions of threads?	🥺 Probably not
It can be used interactively?	🤯 Probably not
It has shiny RGB lights?	...

Transition to next part: UN*X & SSH
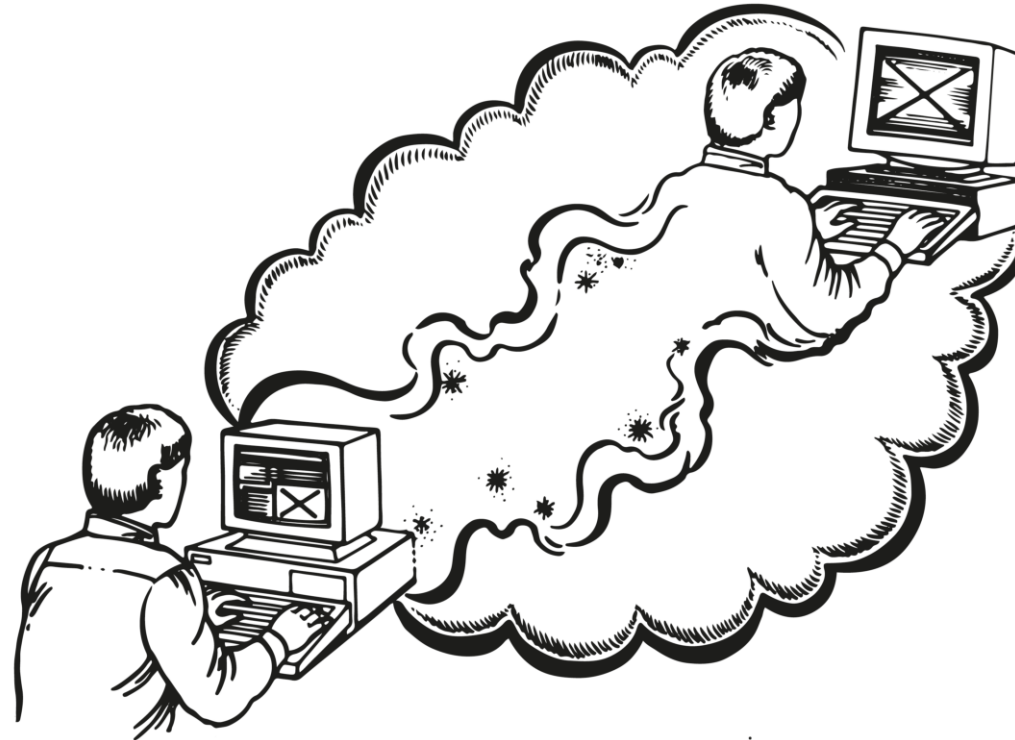
Figure 7.1 Remote login is a lot like astral projection.