High Performance Computing and Modern Science

Prof. Dr. Thomas Ludwig

German Climate Computing Centre Hamburg
Universität Hamburg – Department of Informatics – Scientific Computing

Abstract

High Performance Computing (HPC) became an instrument for many modern sciences. In particular numerical simulations are based on parallel computing in HPC systems. Beside experiment and theory they are considered as the third pillar of knowledge gaining. However, HPC became very expensive over the last ten year. In particular, modern systems ask for huge amounts of electrical power such that the possible research starts to be limited by the research institution´s budget for power. Moreover, programming these systems that exhibit massiv parallelism in their components requires scientists knowledgeable in software engineering. Finally, there is now a new challenge for our researchers: the aggregation of measured and generated digital data leads to new epistemological concepts. It is usually called data intensive science or the Fourth Paradigm. Available data becomes the primary product for new and unique insights. The talk will discuss current challenges in HPC and data intensive science and will highlight the potential development in the field until the end of the decade. Beyond 2020 the development in computer technology is difficult to predict - modern science and modern life ask however for a sustained progress.

Computational Performance

0.123456 x 0.654321

this is ONE floating point operation (FLOP)

if you need 100 seconds to calculate this

your computational performance is 1/100 Flops

Computers and Performance

The first computer in 1941 (Zuse´s Z3) made
0.3 Flops

The No. 1 computer in Nov 2013 (Tianhe-2) makes
33,862,700,000,000,000 Flops
33×10^{15} Flops
33 quadrillion Flops
33 PetaFlops

Computers and Science & Engineering

Zuse´s Z3

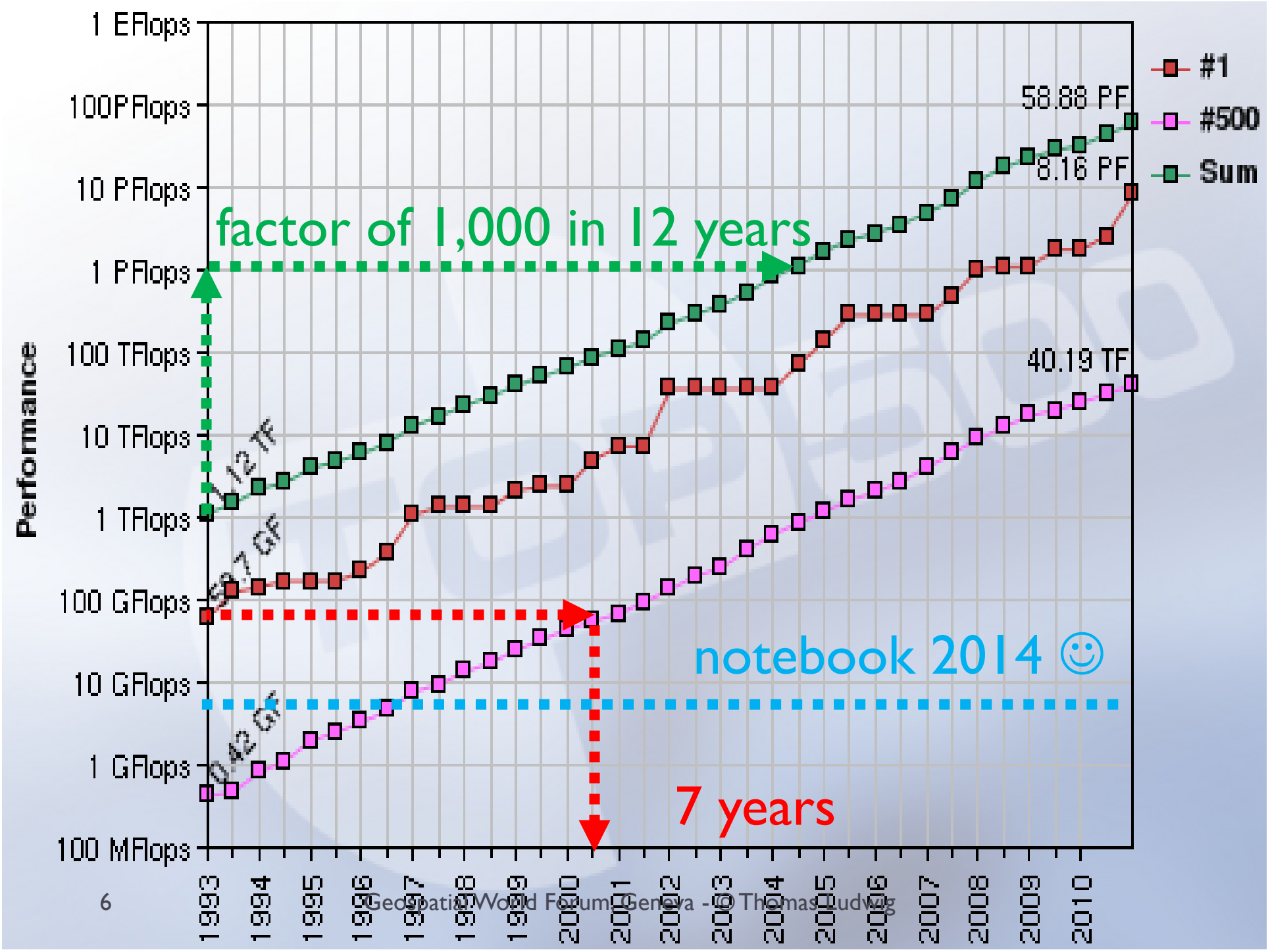
- Fluid dynamics and wing profile optimizations

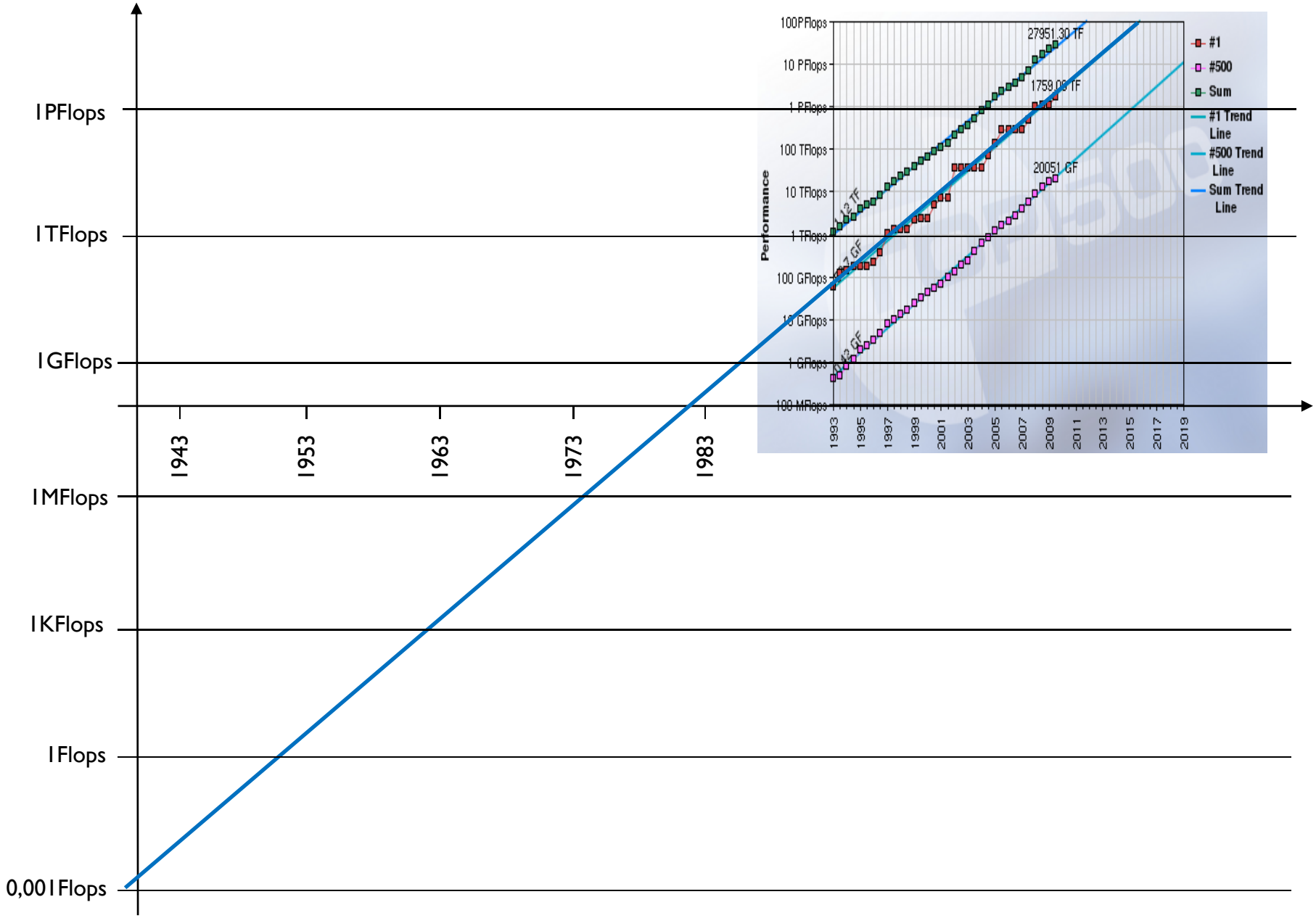
ENIAC (1946, 385 Flops!)

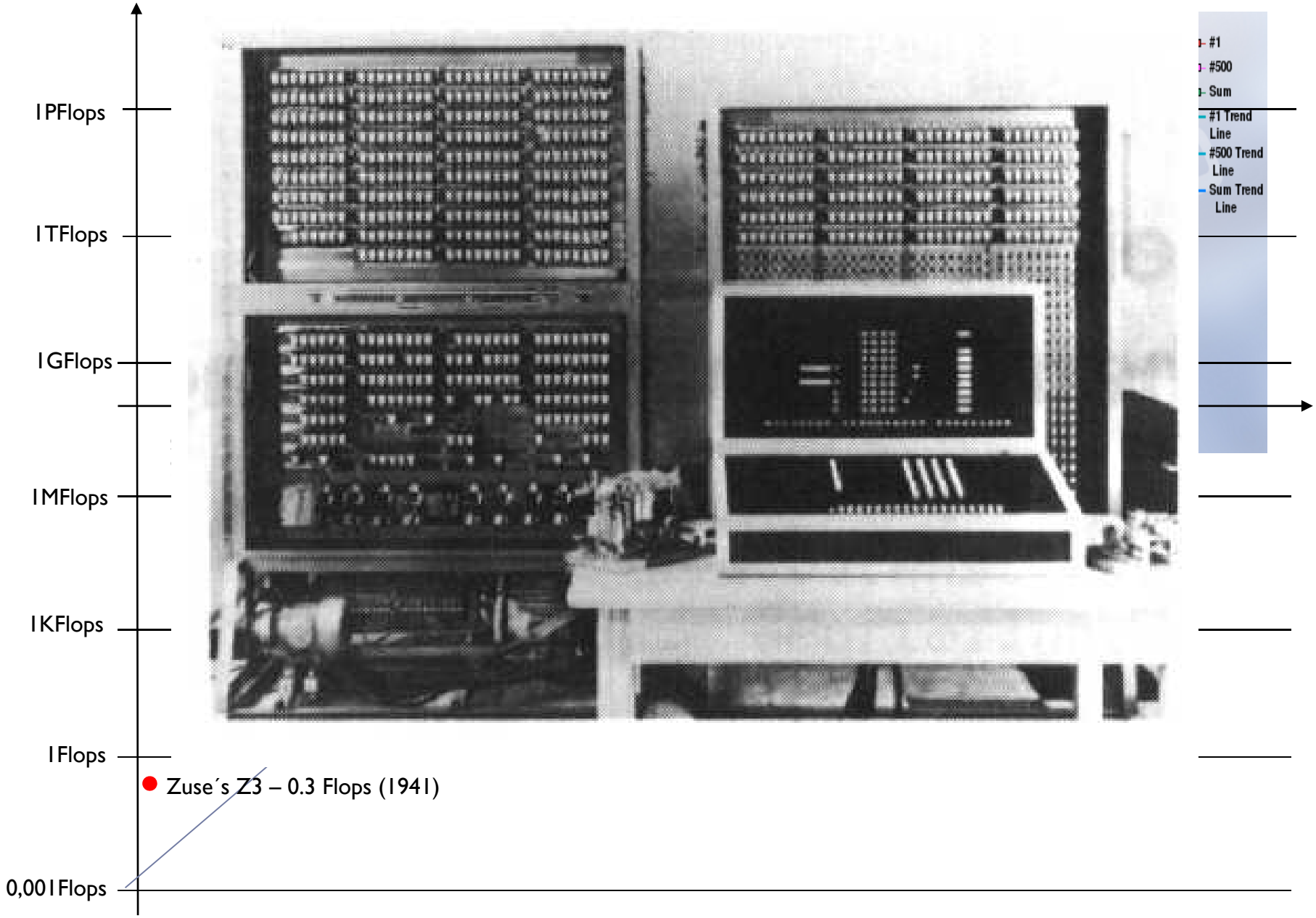
- Artillery firing tables
- Simulate hydrogen bomb (John von Neuman)
- First numerical weather forecast (John von Neuman, 1950)

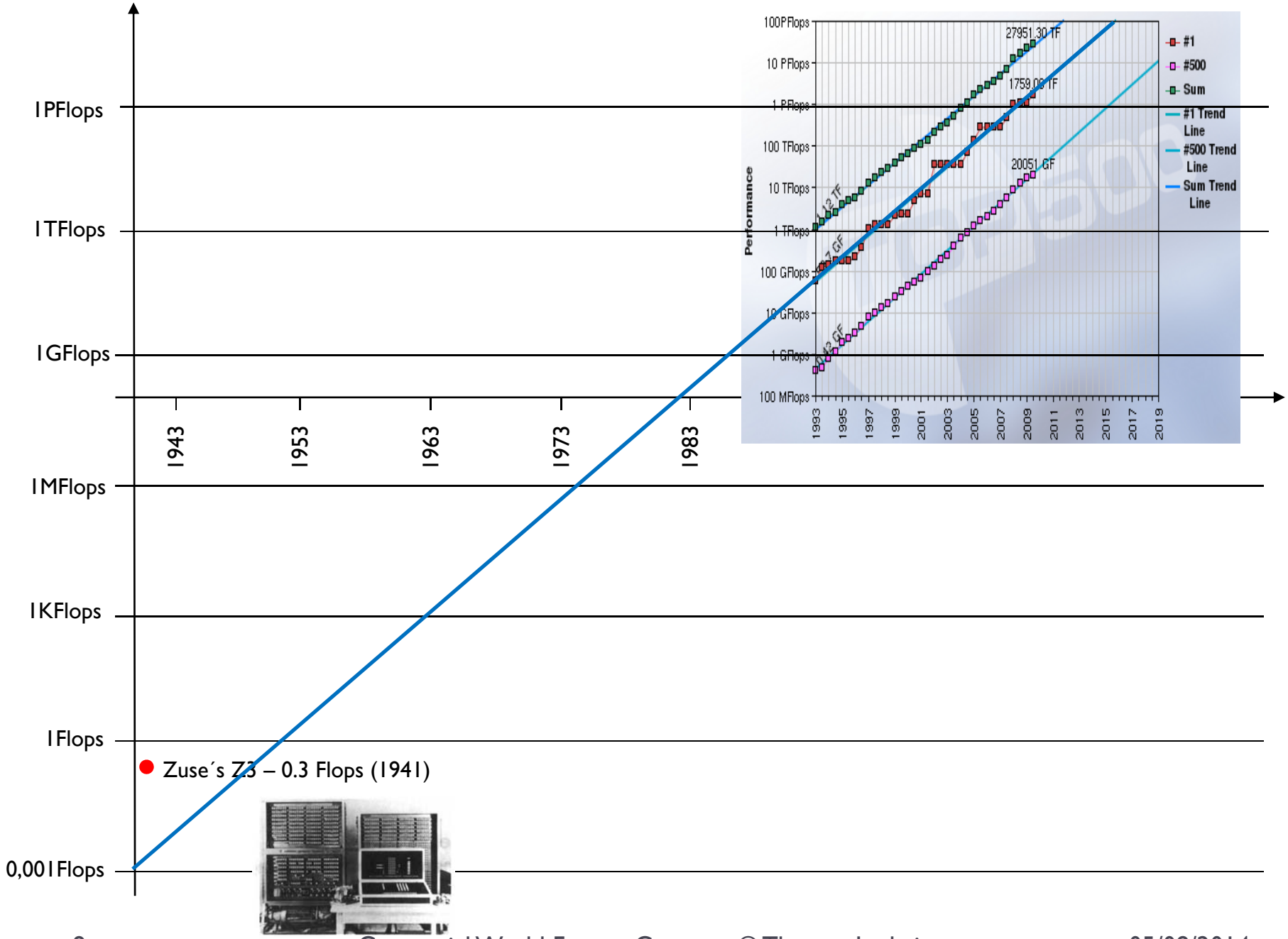
Tianhe-2

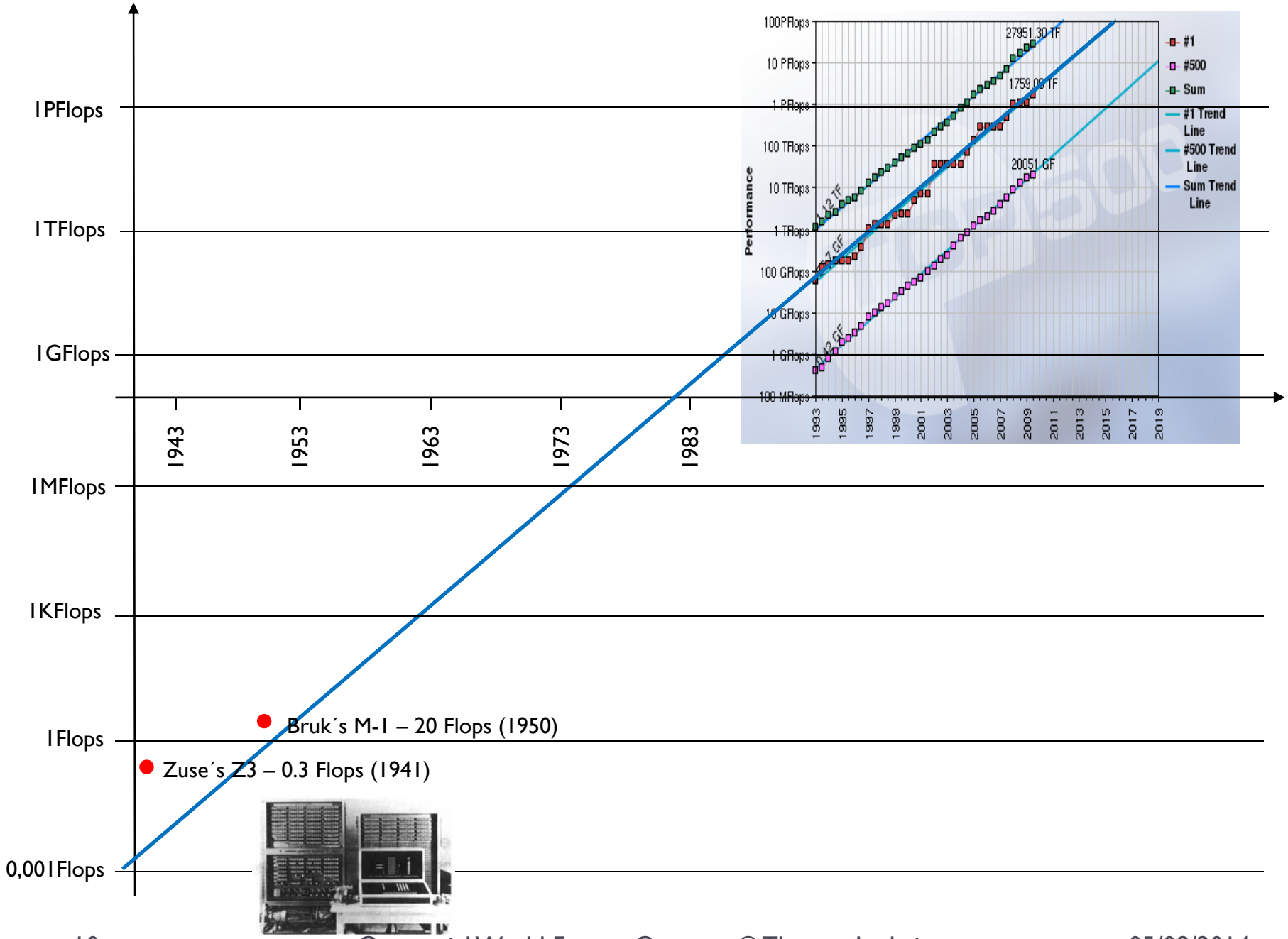
- Fluid dynamics
- Cosmology
- Business Opinion Analysis
- Security e-Government Cloud


1 EFlops
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
Performance
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
#1
#500
Sum
58.88 PF
8.16 PF
40.19 TF
1.12 TF
59.7 GF
0.42 GF
factor of 1,000 in 12 years
notebook 2014 ☺
7 years

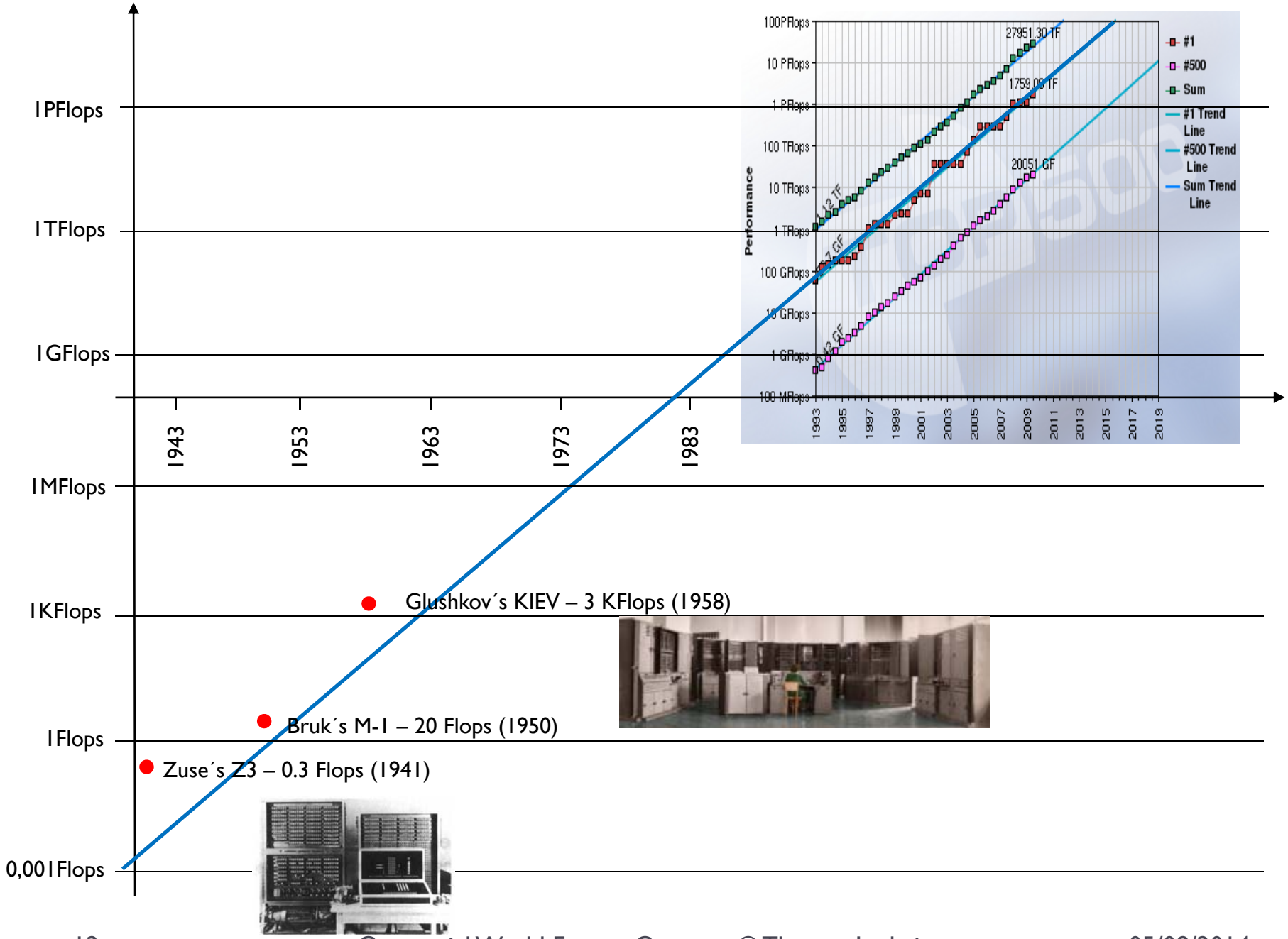

1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
Performance
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
27951.30 TF
1759.09 TF
20051 GF
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line

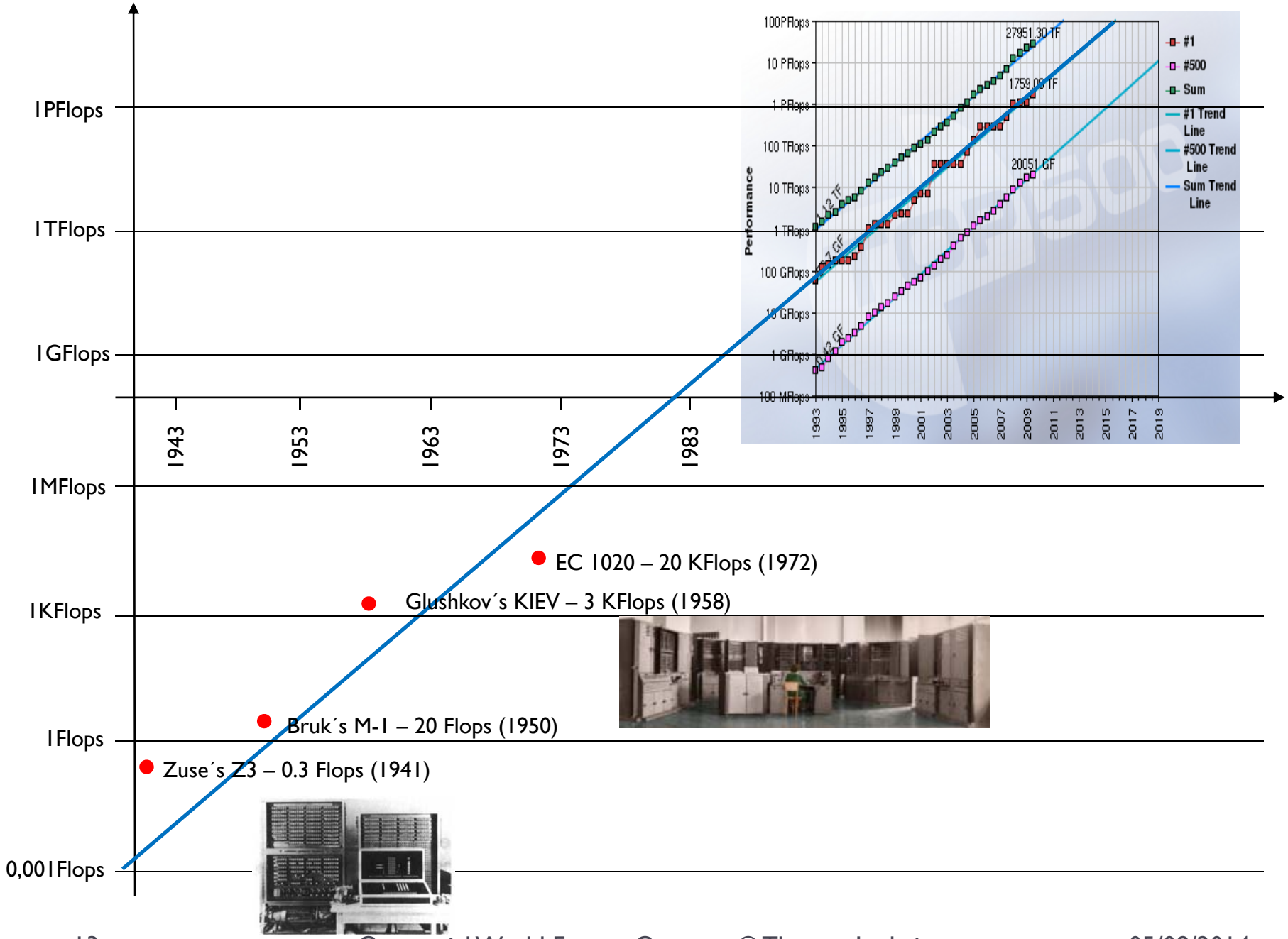

1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
Zuse´s Z3 – 0.3 Flops (1941)
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line

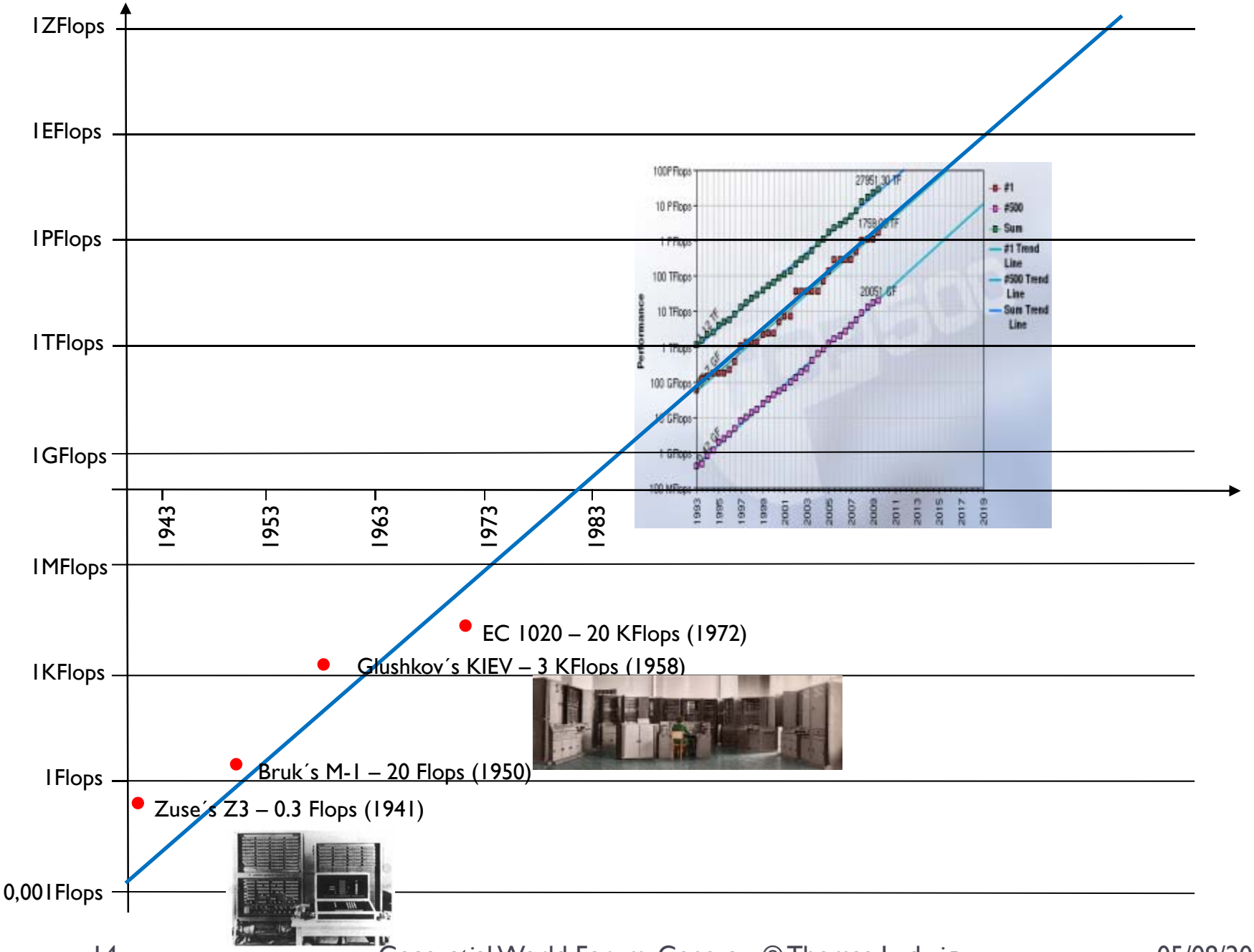

1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
Zuse´s Z3 – 0.3 Flops (1941)
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
Performance
27951.30 TF
1759.09 TF
20051 GF
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019

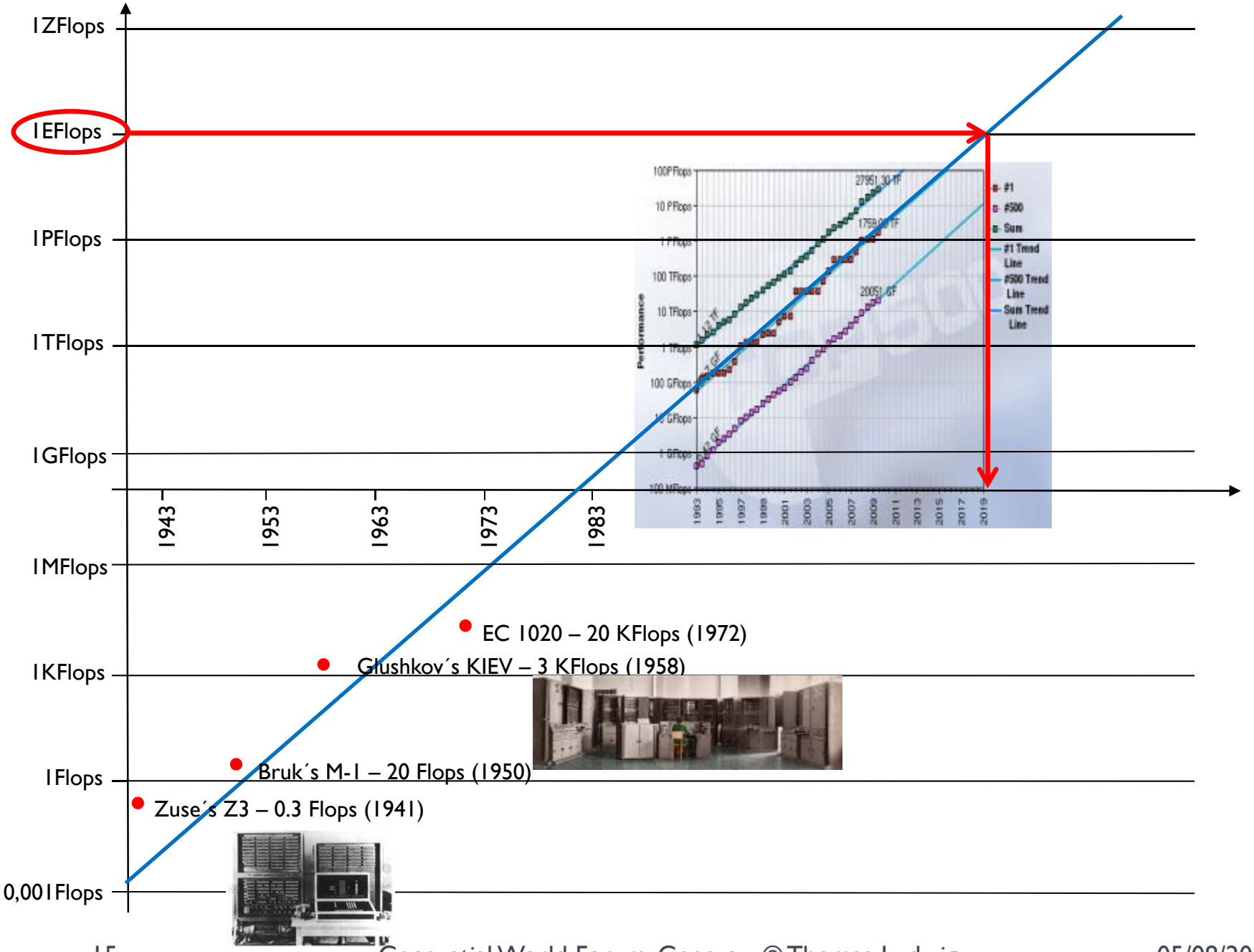

1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
Performance
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
27951.30 TF
1759.09 TF
20051 GF
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line

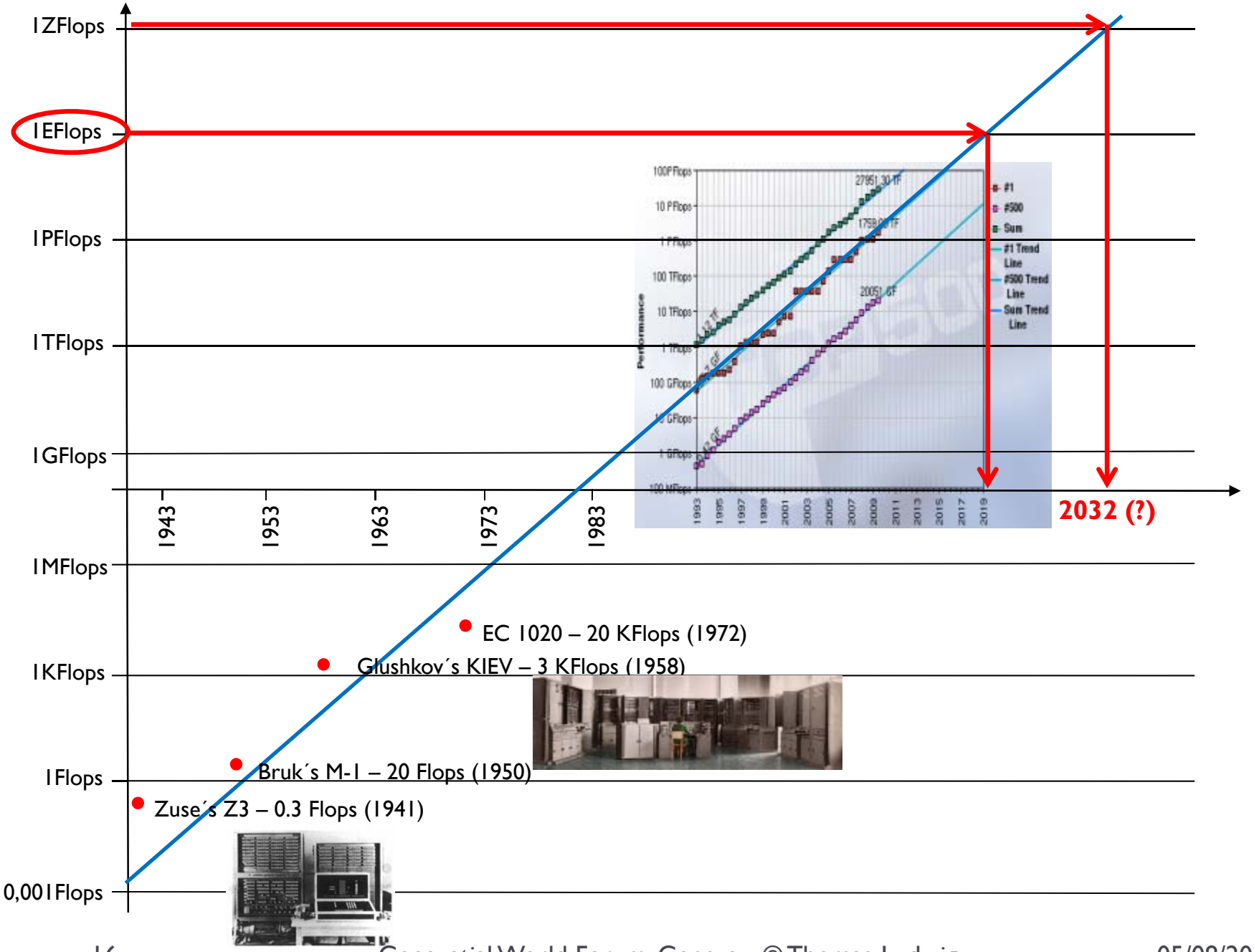
1 PFlops
1 TFlops
1 KFlops
1 Flops
0,001 Flops
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
100 PFlops
10 PFlops
1 PFlops
100 TFlops
27951.30 TF
1759.06 TF
20051 GF
#1
#500
Sum
#1 Trend Line
#500 Trend Line


1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
Performance
27951.30 TF
1759.09 TF
20051 GF
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line


1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
EC 1020 – 20 KFlops (1972)
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
100PFlops
10 PFlops
1 PFlops
100 TFlops
10 TFlops
1 TFlops
100 GFlops
10 GFlops
1 GFlops
100 MFlops
Performance
27951.30 TF
1759.09 TF
20051 GF
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019


1ZFlops
1EFlops
1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
EC 1020 – 20 KFlops (1972)
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
Performance
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line


1ZFlops
1EFlops
1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
EC 1020 – 20 KFlops (1972)
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
Performance
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line


1ZFlops
1EFlops
1PFlops
1TFlops
1GFlops
1MFlops
1KFlops
1Flops
0,001Flops
1943
1953
1963
1973
1983
2032 (?)
EC 1020 – 20 KFlops (1972)
Glushkov´s KIEV – 3 KFlops (1958)
Bruk´s M-1 – 20 Flops (1950)
Zuse´s Z3 – 0.3 Flops (1941)
Performance
#1
#500
Sum
#1 Trend Line
#500 Trend Line
Sum Trend Line

Increase in Computational Performance

from 0.3 Flops in 1941

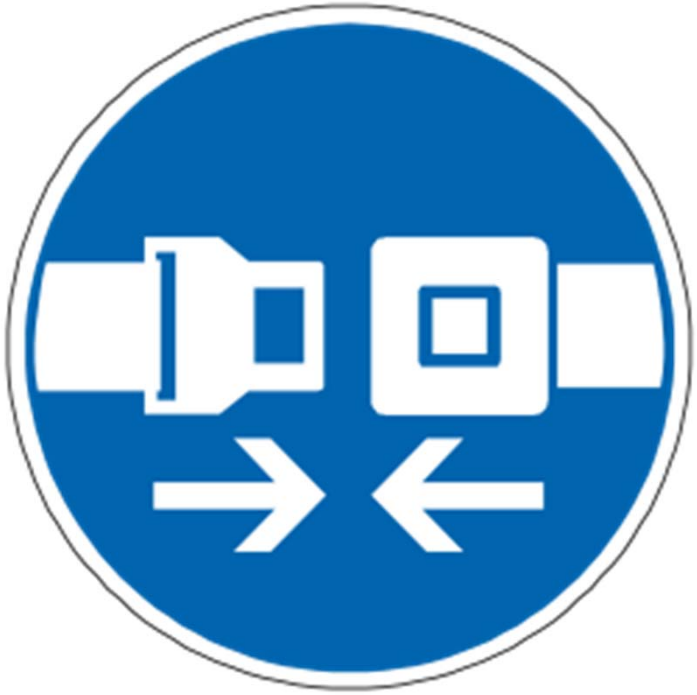
to 33×10^{15} Flops in 2013

as constant exponential growth

means factor 1.000 every 12.5 years

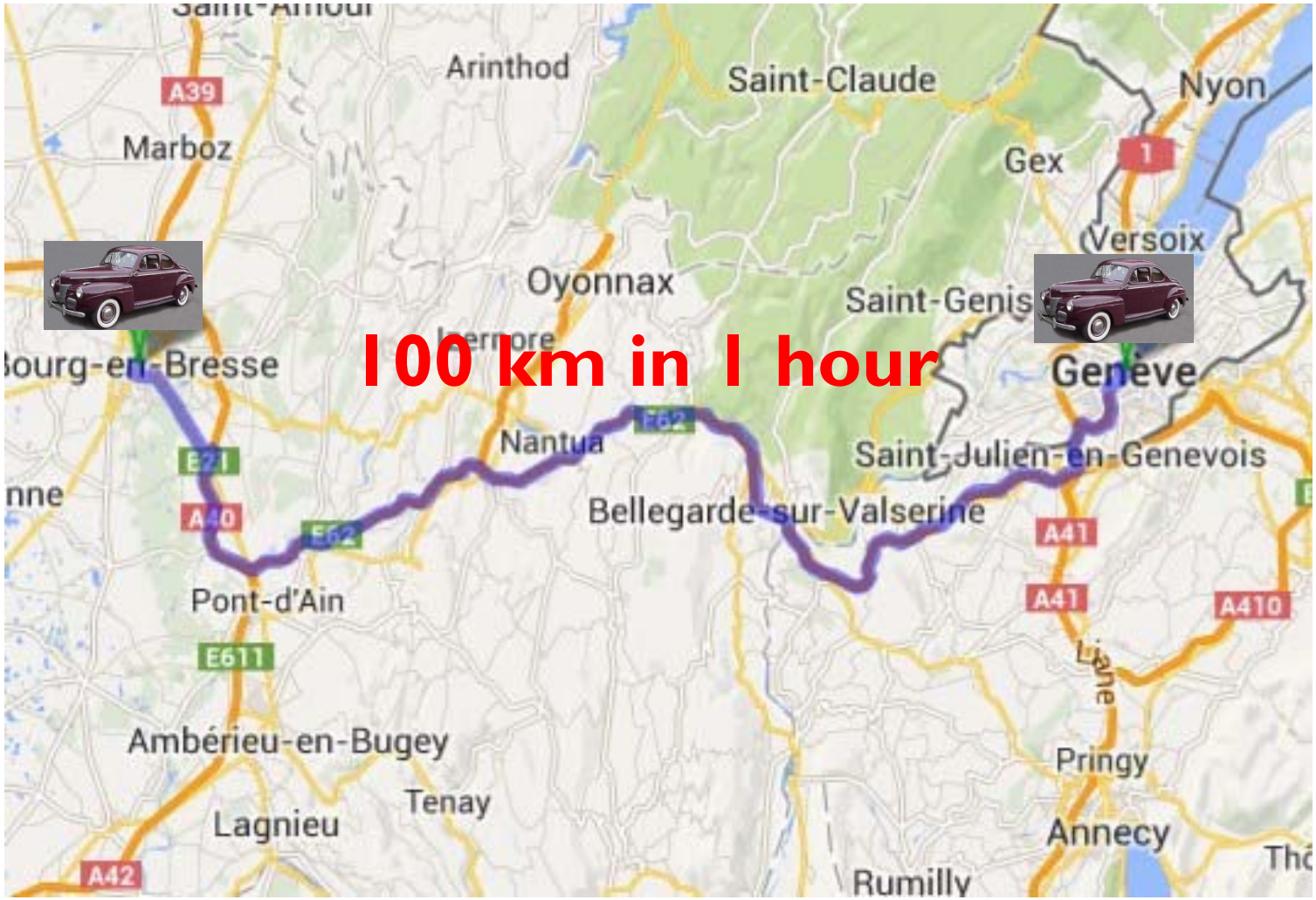
for already more than 70 years

Exponential Acceleration

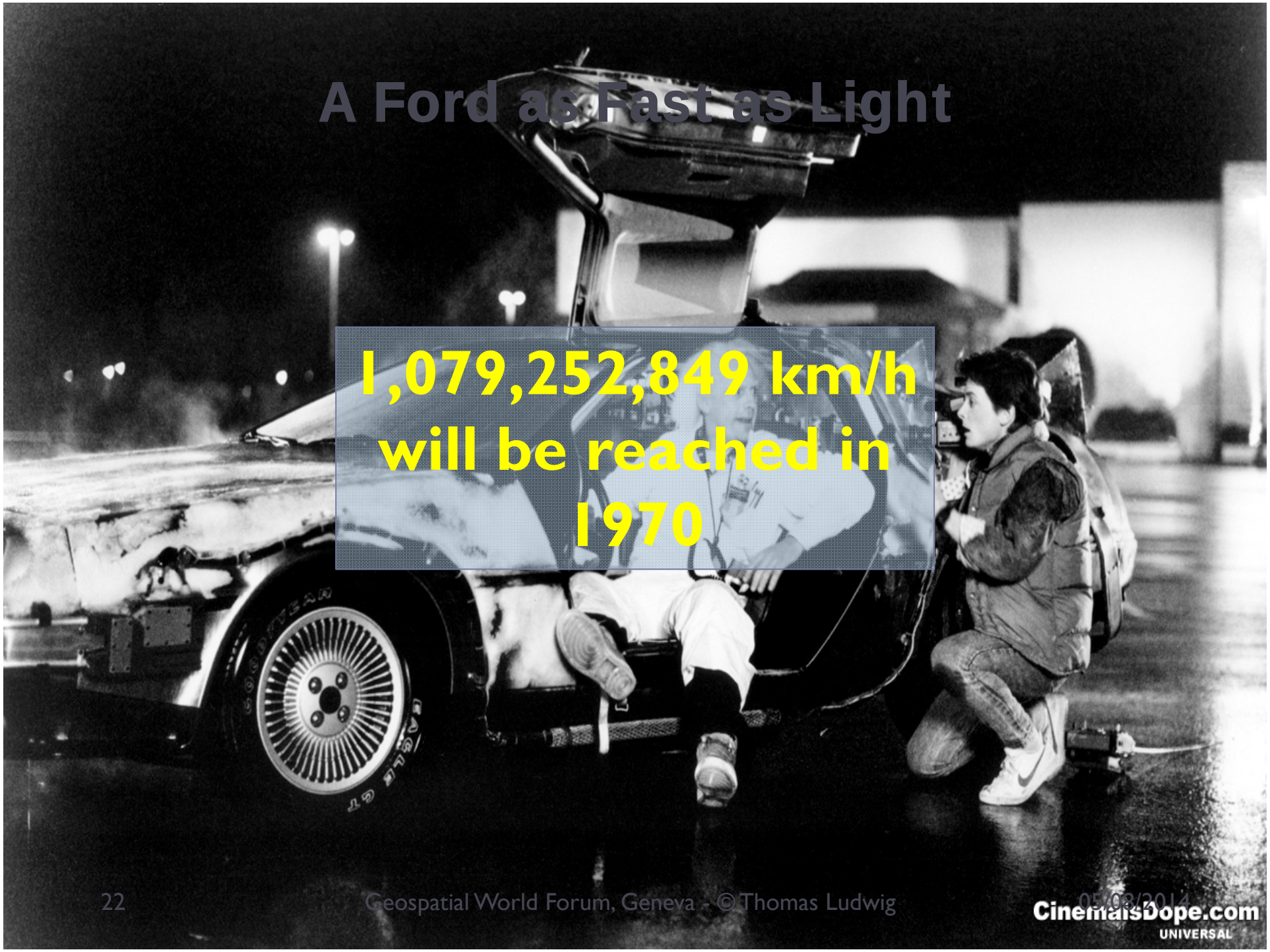


A 1941 Ford

Geneva – Bourg-en-Bresse

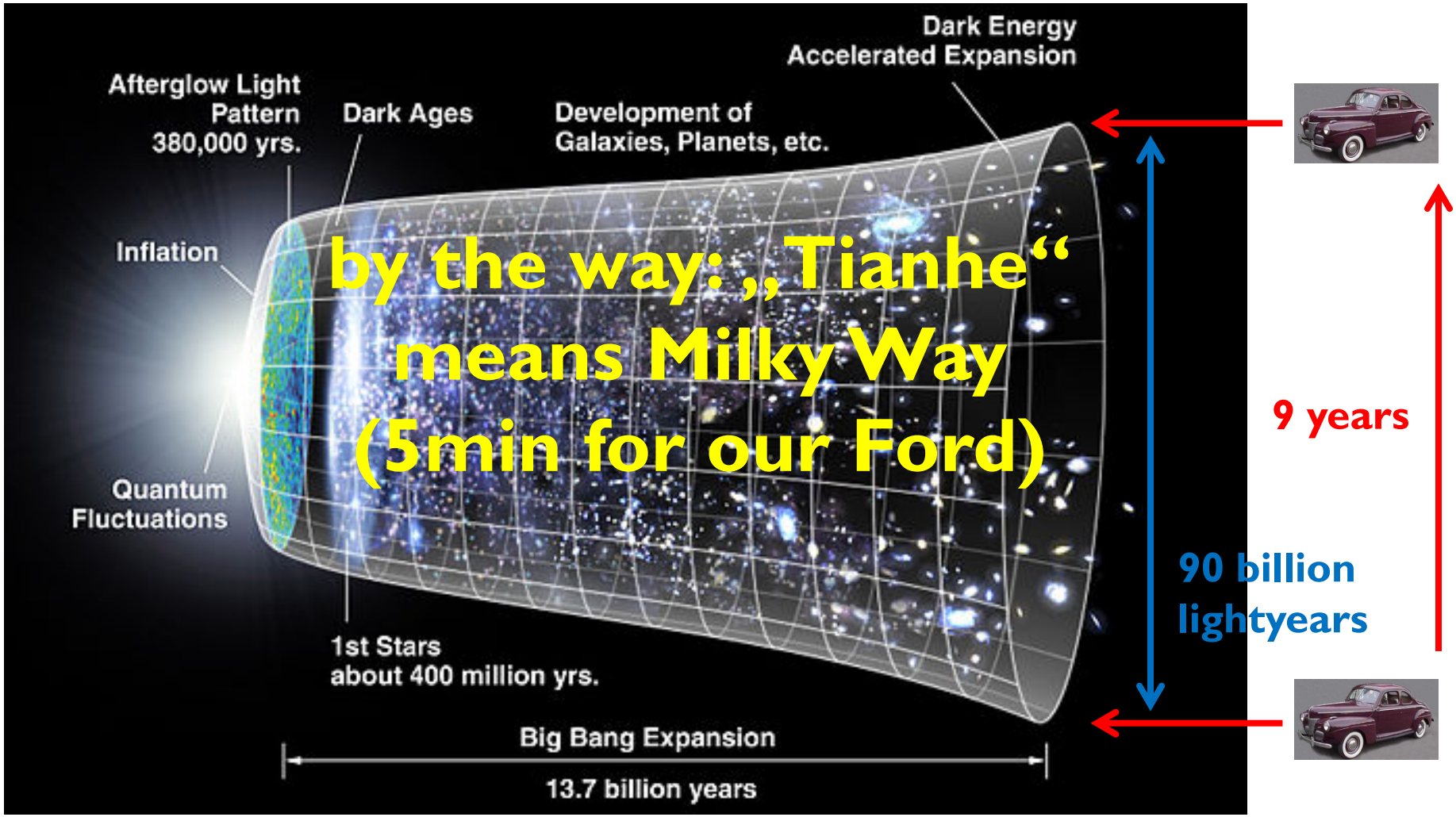


A Ford as Fast as Light

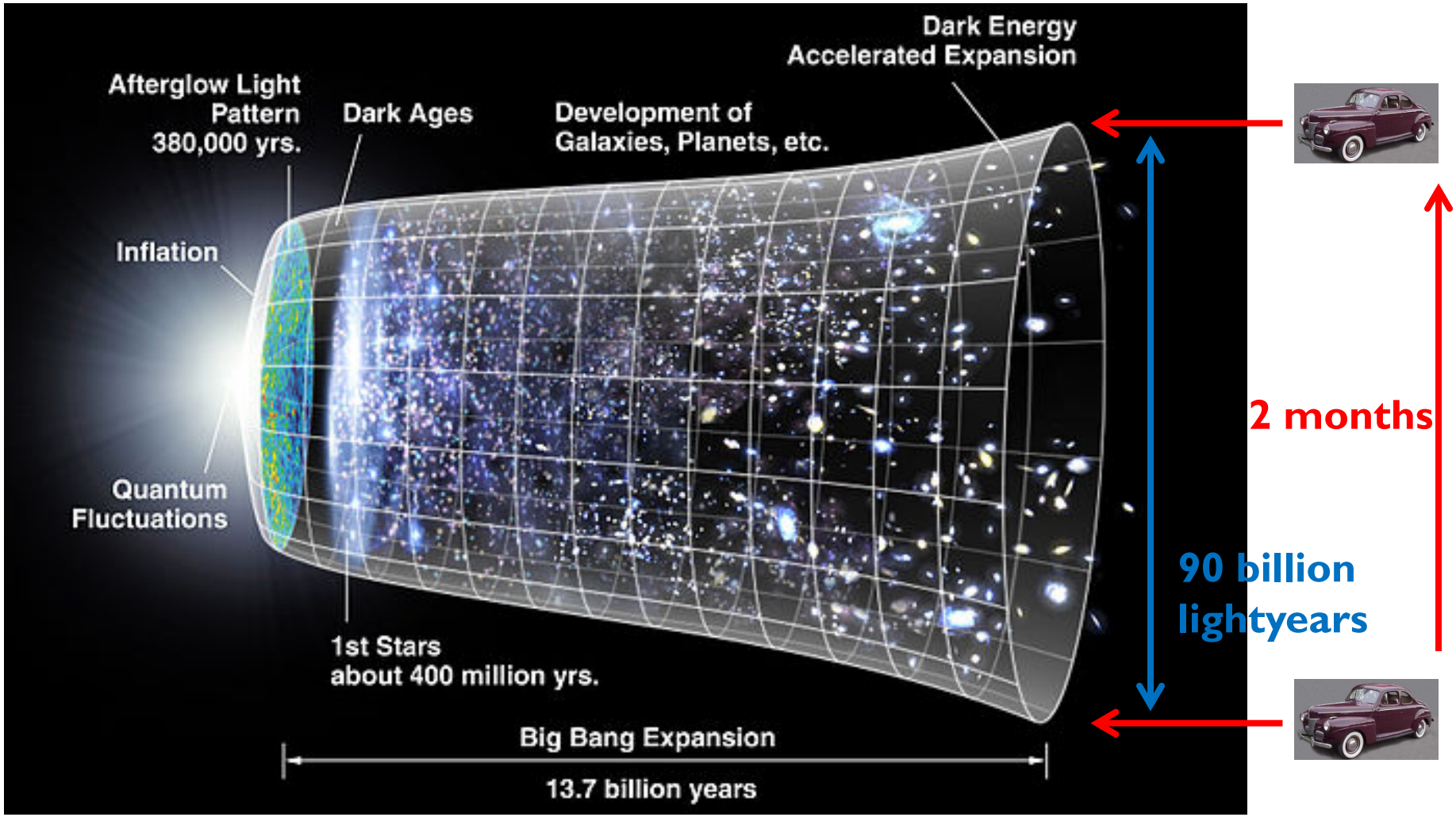


1,079,252,849 km/h will be reached in 1970

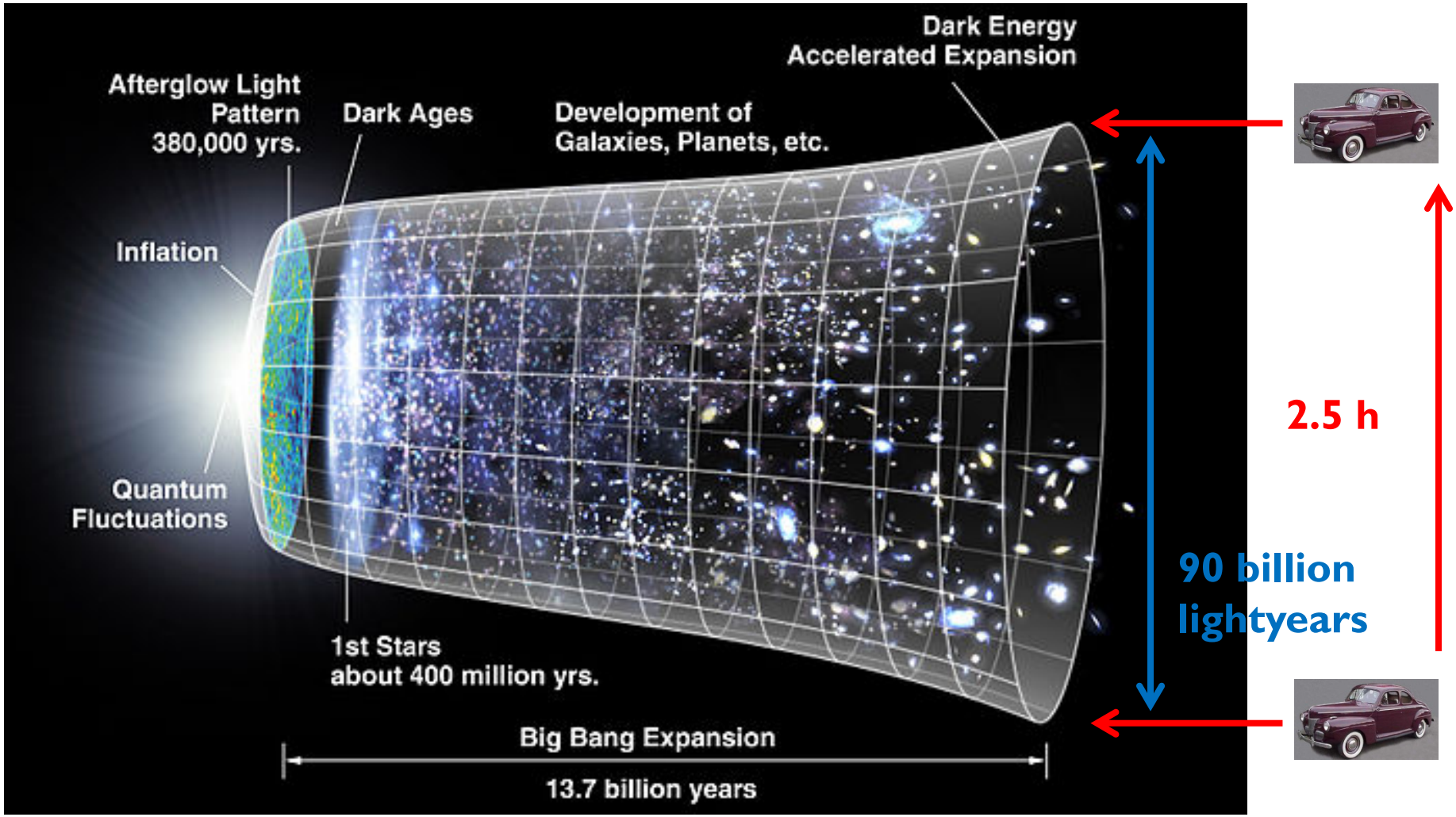
The Universe and the 2013 Ford (Petascale)



The Universe and the 2020 Ford (Exascale)



The Universe and the 2032 Ford (Zetascale)



Energy Consumption

A killer argument against
high performance computing today and tomorrow

- Tianhe-2 uses 17 MW
- May cost 17 million Euro per year for power

- Same problem with most compute HPC centers

But…

Time will tell...

Currently energy efficiency increases
by a factor of 10 every 7 years

1987

25 years of DKRZ

2012

What Type of Science?

Basically we do numerical simulations

- Climate and weather simulation
  - Cloud simulation
- Cosmology
  - The origin of the universe
- Particle physics
- Proteine folding
  - Individualized medical treatment
- Brain simulation

Orders of Magnitude

Human brain

- 100×10^9 neurons

Tianhe-2 in 2013

- 33×10^{15} Flops
  - 330.000 Flops/neuron

Exaflops in 2020

- 1×10^{18} Flops
  - 10 million Flops/neuron

Zetaflops in 2032

- 1×10^{21} Flops
  - 10 billion Flops/neuron

dataddeluge

The Fourth Paradigm

Refer to Jim Gray et al.:

"The Fourth Paradigm –
Data-Intensive Scientific Discovery"

„Data are not consumed by the ideas and innovations they spark but are an endless fuel for creativity"

„Harnessing the Power of Digital Data for Science and Society"
by the Interagency Working Group on Digital Data
January 2009

Climate Science and Data

Climate research was and is dependent on data

- Keep data for long term in order to validate models in the future
- Keep all model variables because we do not know what will be necessary in future analyses
- Use a fine global grid in order to select later local regions
- ...

DKRZ archived around 8 PB on tape in 2013

- This is all output data of simulation runs

Data Generation

DKRZ (115 TFlop/s, 26 TByte main memory) produces an estimated data transfer mem<->disk

- 5-10 GB/s (430-860 TB/day)
- ca. 100 TB/day are saved for further inspection
- ca. 20 TB/day are archived to tape

Example CMIP5

The 5th Coupled Model Intercomparison Project

- Provides key input for the next IPCC report (5th AR)
  - Intergovernmental Panel on Climate Change
- ~20 modeling centers around the world (DKRZ being one of the biggest)
- Produces 10s of PBytes of output data from ~60 experiments ("digital born data")

Data are produced without knowing all applications beforehand and these data are stored and archived for interdisciplinary utilization by yet unknown researchers

Example Workflow CMIP5

Data creation

- German contribution produced at DKRZ
- Used ½ year of its compute capacity

Data archival

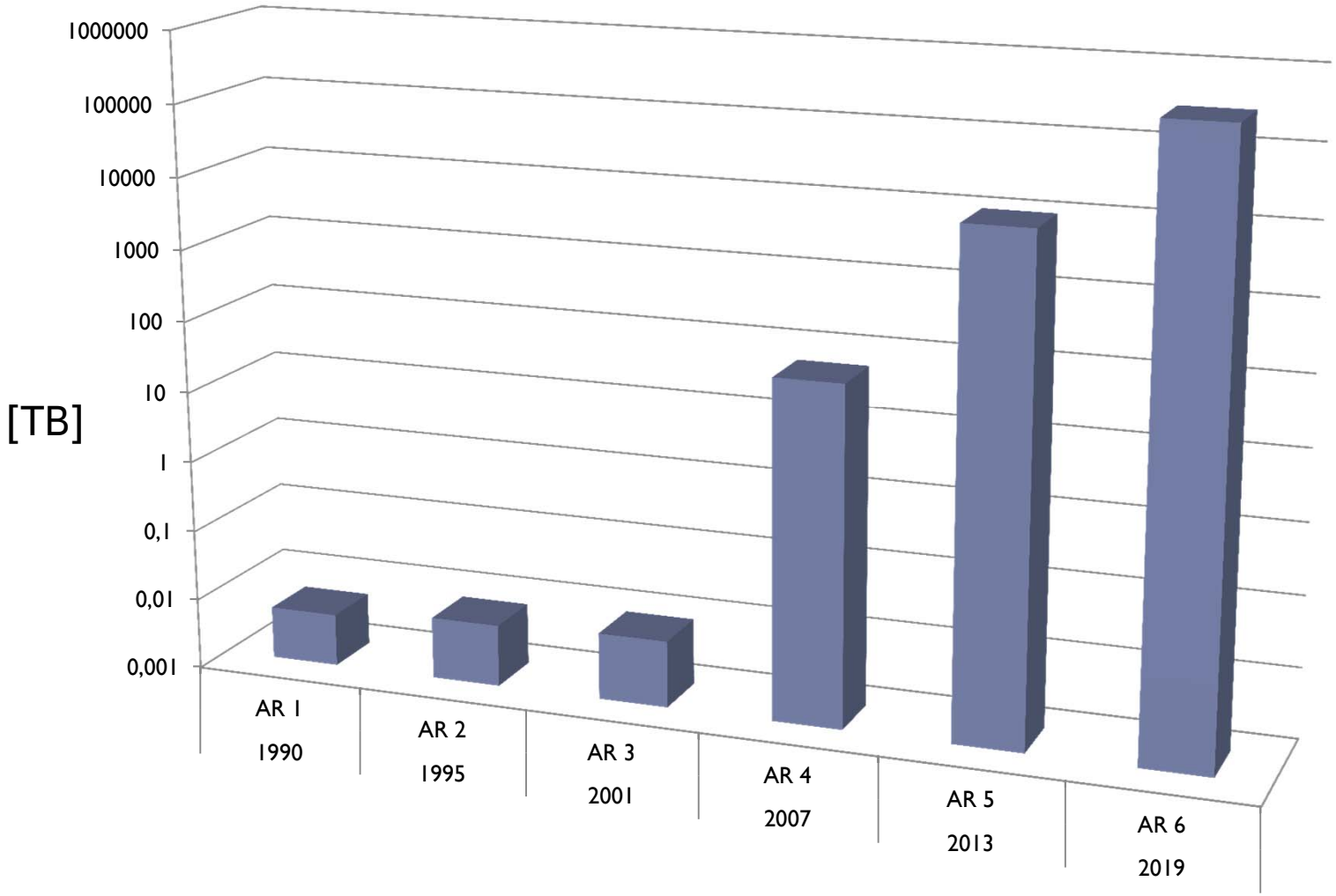
- 600 TB raw + 60 TB with quality control

Data dissemination

- Will be heavily used during the next 5 years

Expected CMIP6 data volume: factor 20-50 higher

The Next 10 Years



Future Computer and Storage

Next generation climate computer

- 2-3 PFlop/s
- 45 PByte on disk
- An estimated ½ EByte as tape library capacity

- Storage: 30-50% of investment and energy costs
- Currently it is more like 10%

- Storage 50% of the overall complexity ?

Where will this leads us to?

Where will this lead us to?

- Unprecedented science
  - E.g. life sciences: genetic design, articificial DNA
- Unprecedented products
  - E.g. medical science: prosthesis, medical treatments

- Unprecedented problems
  - Ethical question will be undiscussed
  - Societal considerations will be undiscussed
  - Legal aspects will be undiscussed

Technological Singularity

The technological singularity, or simply the singularity, is a hypothetical moment in time when artificial intelligence will have progressed to the point of a greater-than-human intelligence, radically changing civilization, and perhaps human nature. Because the capabilities of such an intelligence may be difficult for a human to comprehend, the technological singularity is often seen as an occurrence beyond which the future course of human history is unpredictable or even unfathomable. (Wikipedia)

"Singularity" – mentioned first in 1958 by John von Neumann
See also his book *The Computer and the Brain*

See Ray Kurzweil´s non-fiction book: "The Singularity is Near"