Evolution of computational infrastructure over the next decade

ET-ILIAS_GWA joint meeting
24-26 Nov 2008

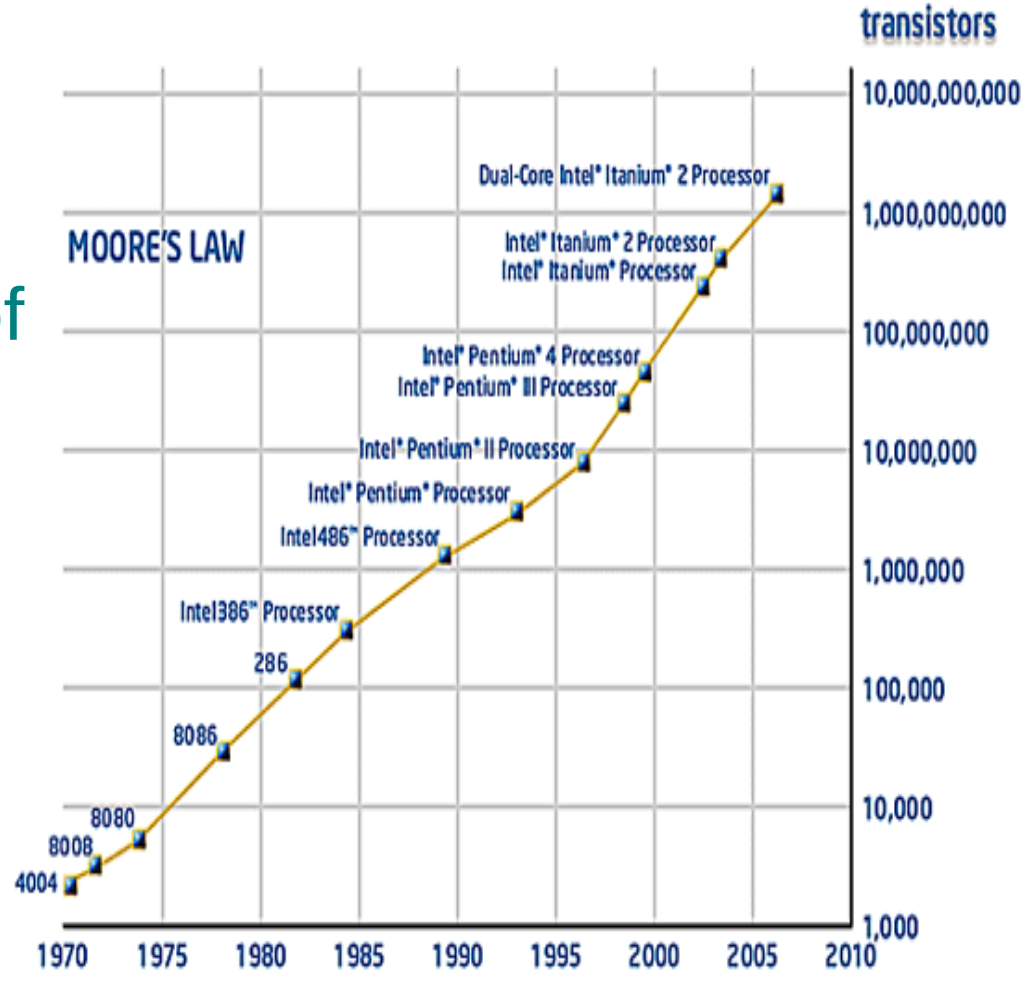
Dr. Leone B. Bosi
INFN and University of Perugia (Italy)

Introduction

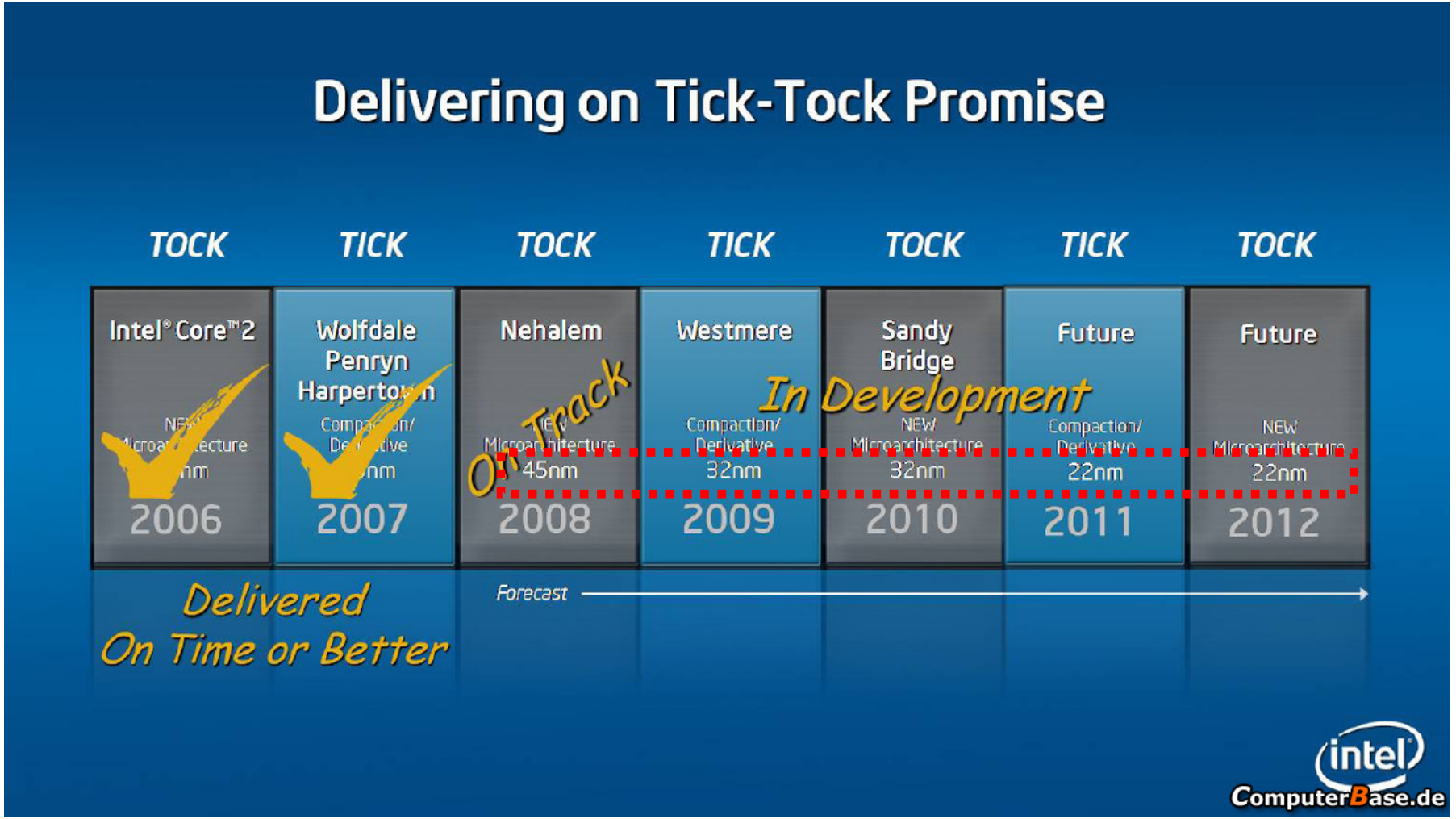
- Computing has made great strides in recent years.
- But as much as it has advanced in the last ten years will advance in the next decade
  - demands for much higher performance, for much lower power density and for greatly expanded functionality
- Continuing increases in transistor count are vital but the next step is also going to take some rethinking of basic foundations such as:
  - Process technology
  - Architecture
  - Software

Moore's Law

- Moore's Law: "Since the invention of the integrated circuit in 1958, the number of transistors on a given chip double every two years"
- This is an exponential growth that has allowed computers to get both cheaper and more powerful at the same time.



Intel roadmap up to 2012

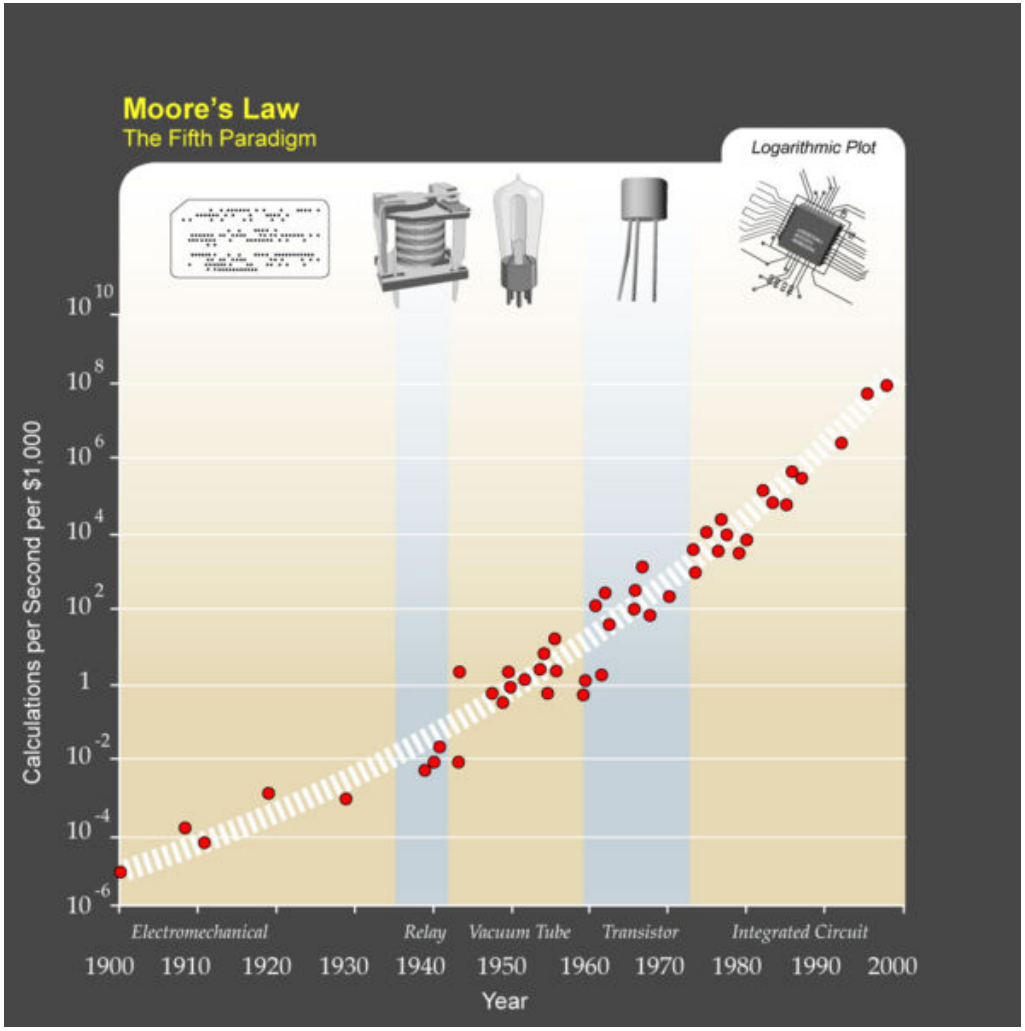


Will computers reach top speed by 2020?

- Moore's Law, as chip manufacturers generally refer to it today, **is coming to an end**, according to a recent research papers.
- Mr Moore said: "**by about 2020, my law would come up against a rather intractable stumbling block: the laws of physics**." The end of the Moore's Law?
- The era in which computing power doubles every toughly 2 years is drawing to a close.
- In **2018** (roughly) Manufacturers will be able to produce chips on the **16-nanometer** manufacturing process, and maybe one or two manufacturing processes after that, **but that's it.**

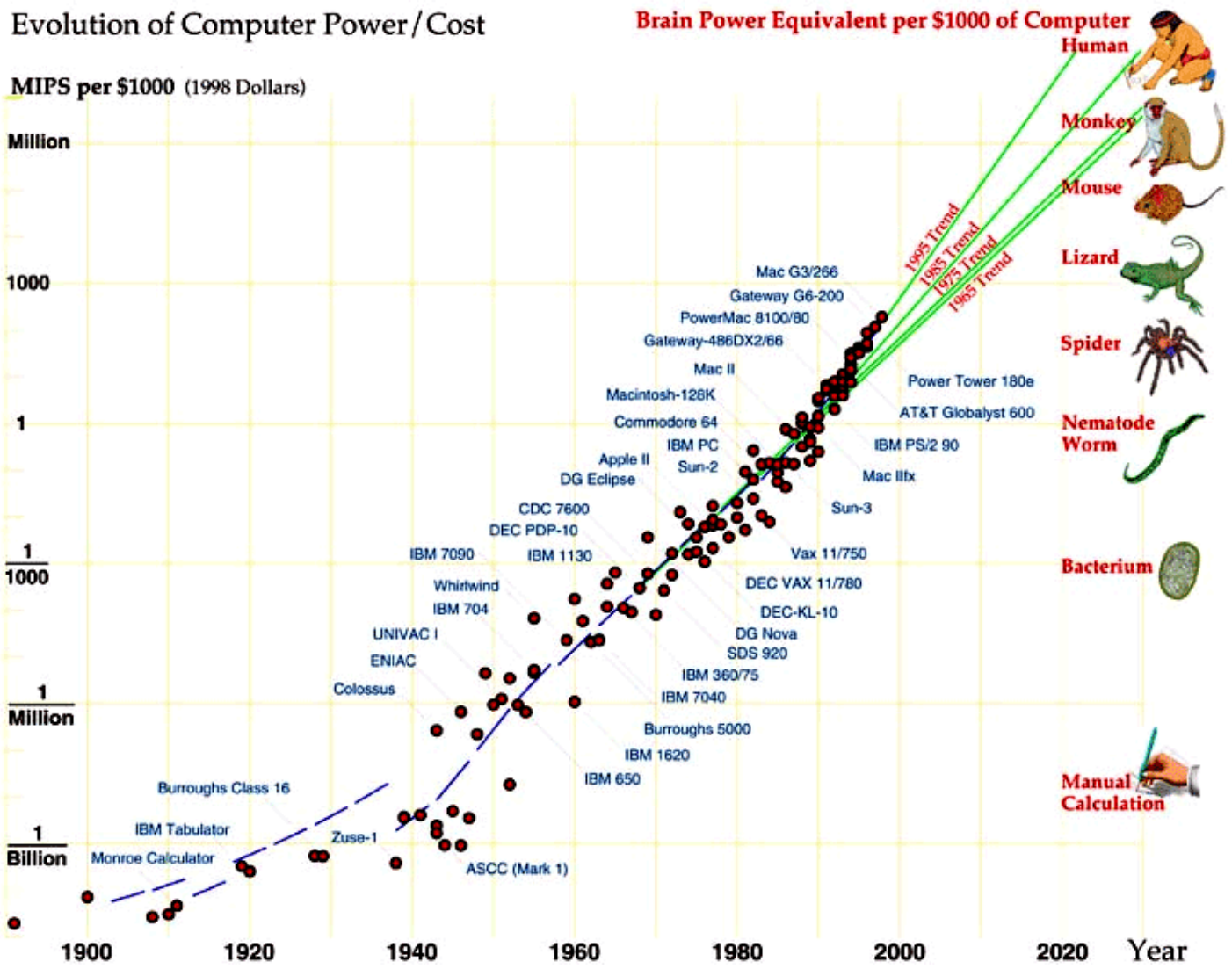
Will computers reach top speed by 2020?

- A possible analysis made by of the history of technology shows that technological change is exponential, contrary to the common-sense "intuitive linear" view.
- E.g.Kurzweil conjectures
  - The Moore's low is only the fifth paradigm
  - Some new type of technology will replace current integrated-circuit technology
- We can find some new phenomenological behaviors permitting to hold true the Moore's Law long after 2020!

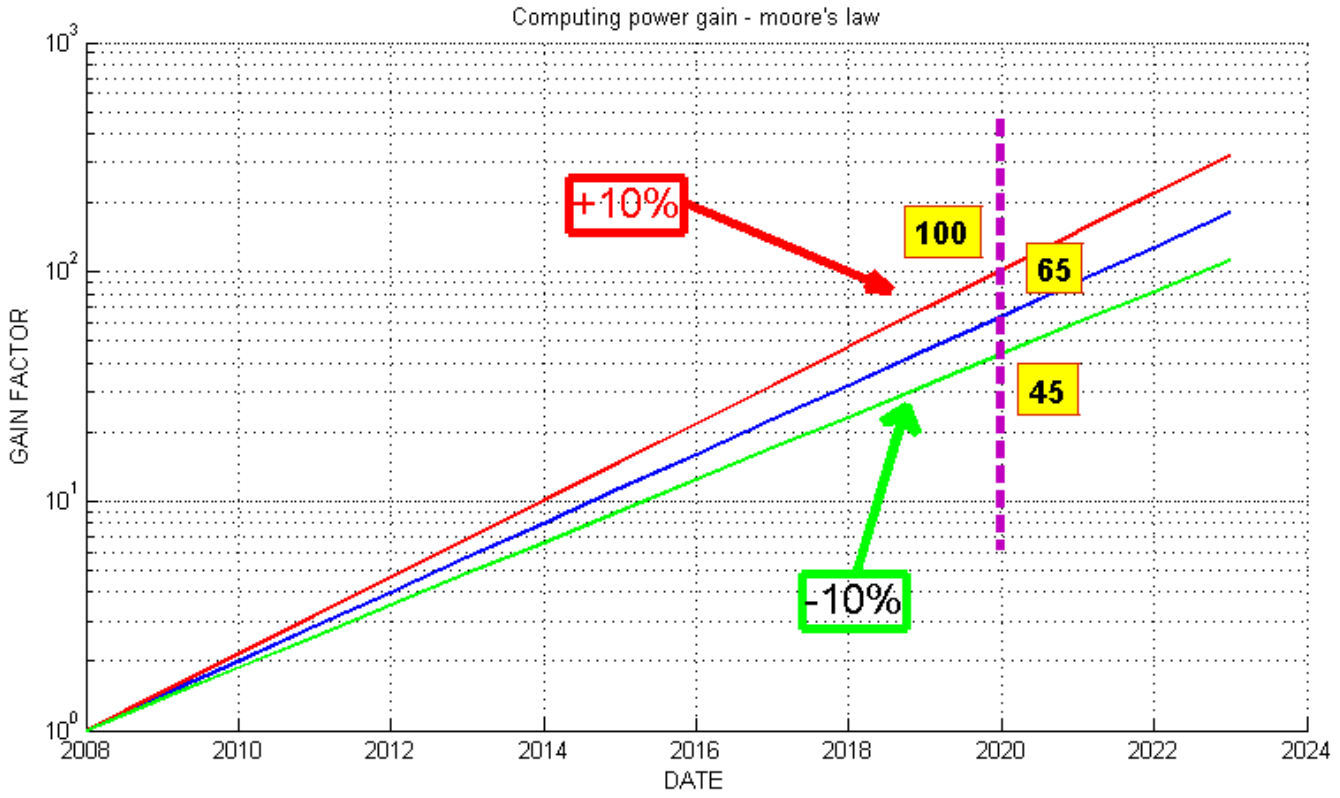


REF: http://www.kurzweilai.net/articles/art0134.html?printable=1

Playing with the future: 2020 forecast



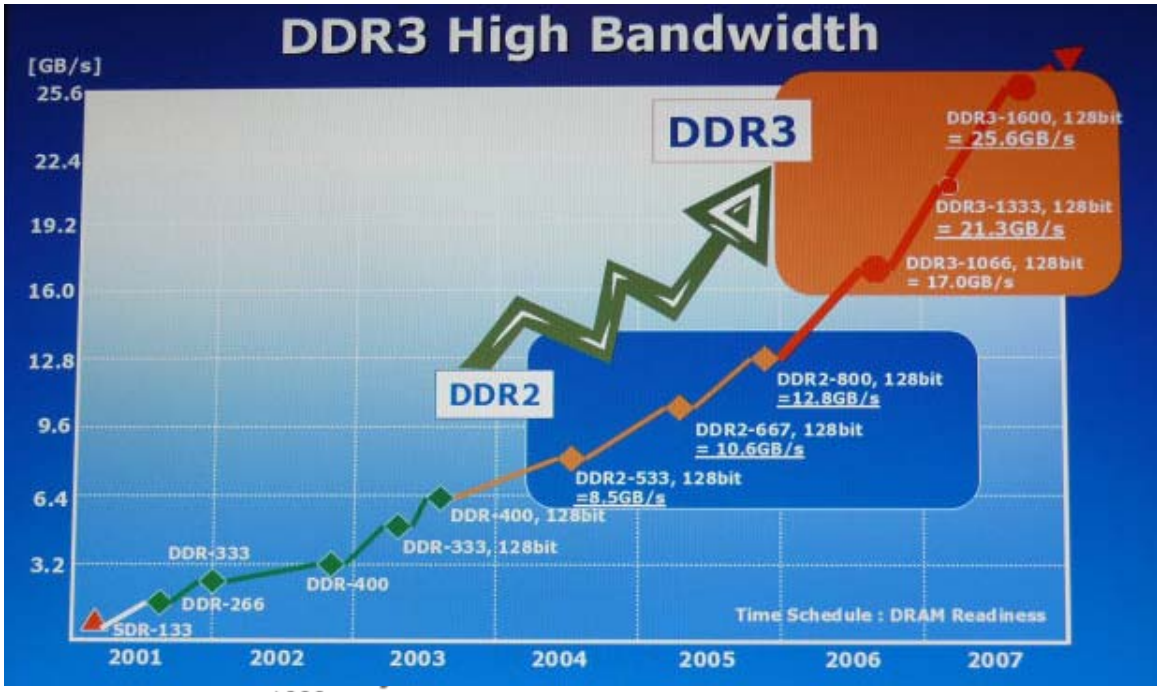
Computing power forecast for 2020
A conservative scenario

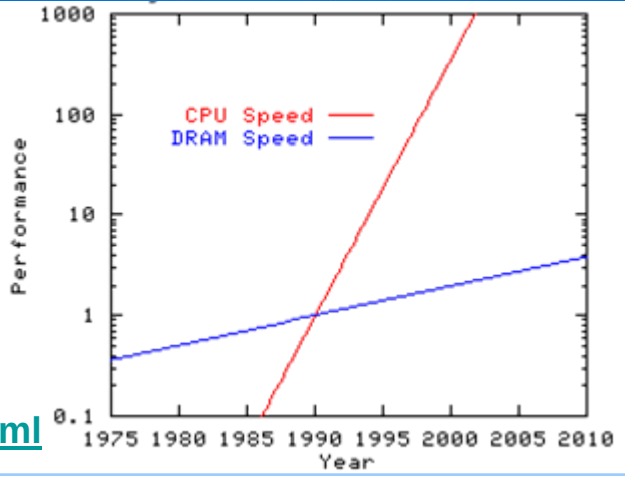


- We can consider a wide range of possible scenarios considering an initial error
- We can venture a conservative forecast about the computing power **gain for 2020-2022 of roughly 100 respect to now.**

Memory technology evolution

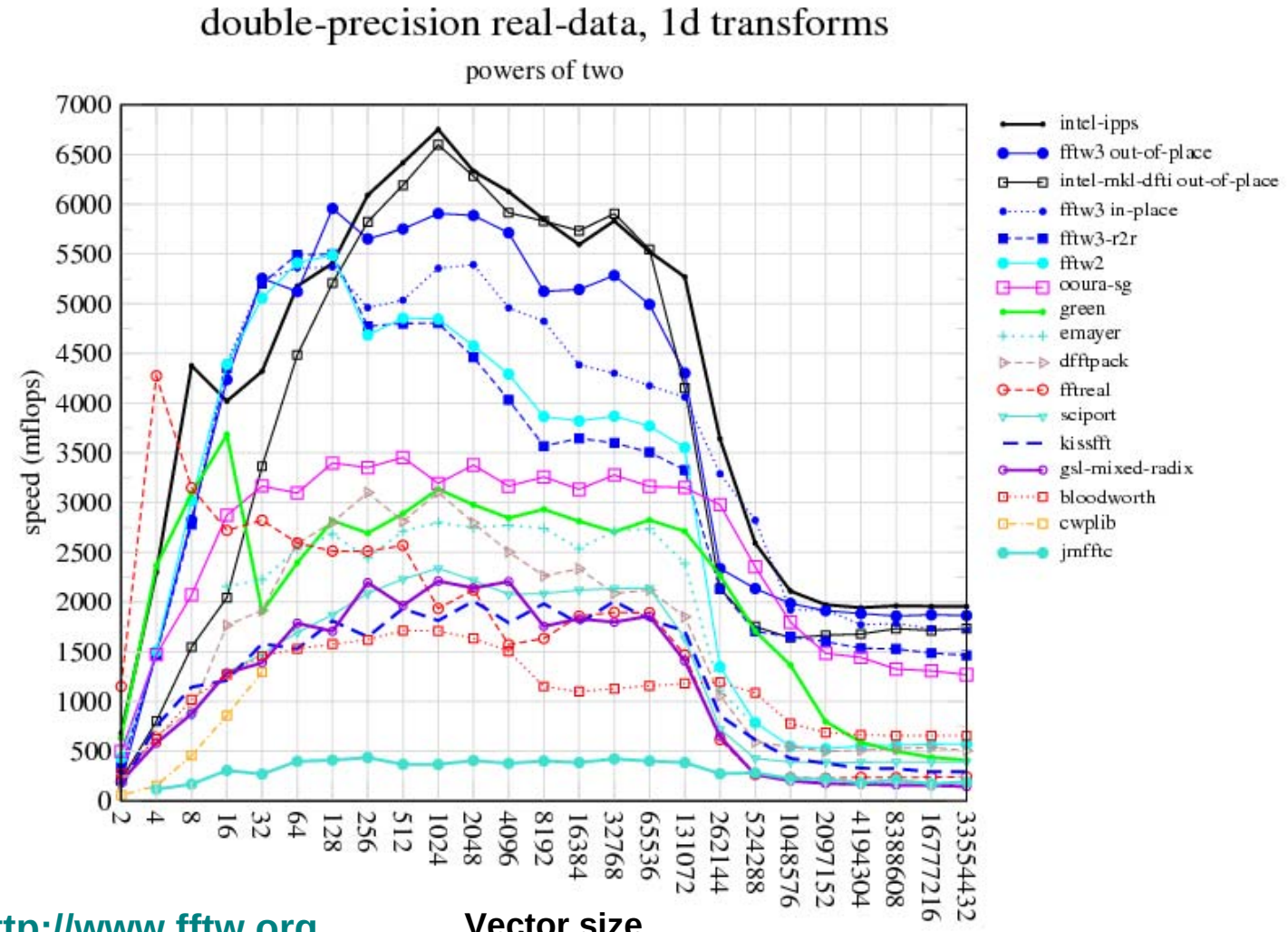
- Memory evolution speed is slower that the CPU evolution speed
- This means that for high Input/Output demanding problems the memory IO is the main bottleneck and not the CPU power
- A well known example of it is the FFT algorithm





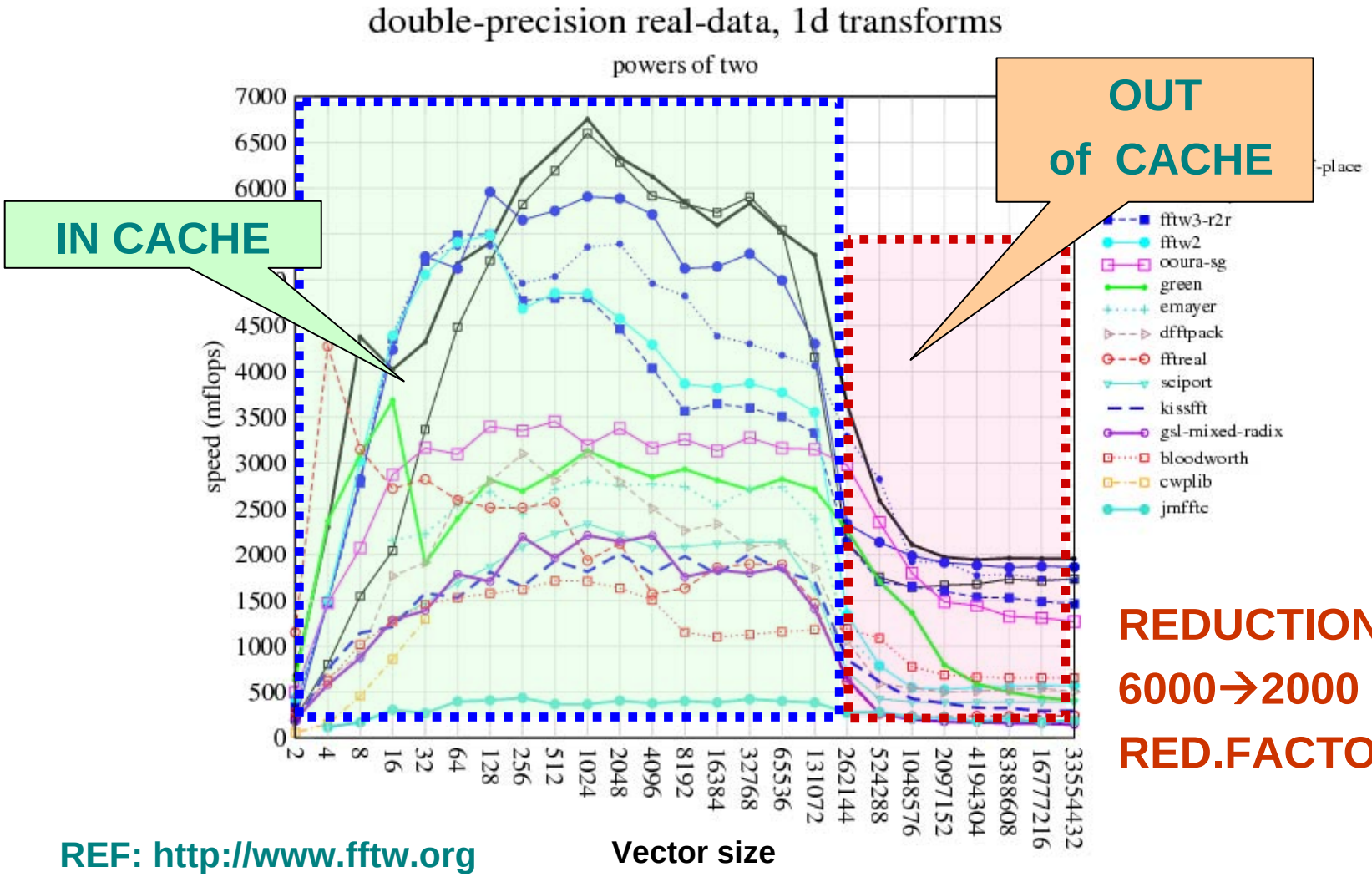
REF: Intel ; http://www.cs.virginia.edu/stream/ref.html

FFT benchmark
3.0 GHz Intel Core Duo, Intel compilers, 64-bit mode



REF: http://www.fftw.org

FFT benchmark
3.0 GHz Intel Core Duo, Intel compilers, 64-bit mode



REDUCTION:
6000→2000 MFLOPS
RED.FACTOR= 3

REF: http://www.fftw.org

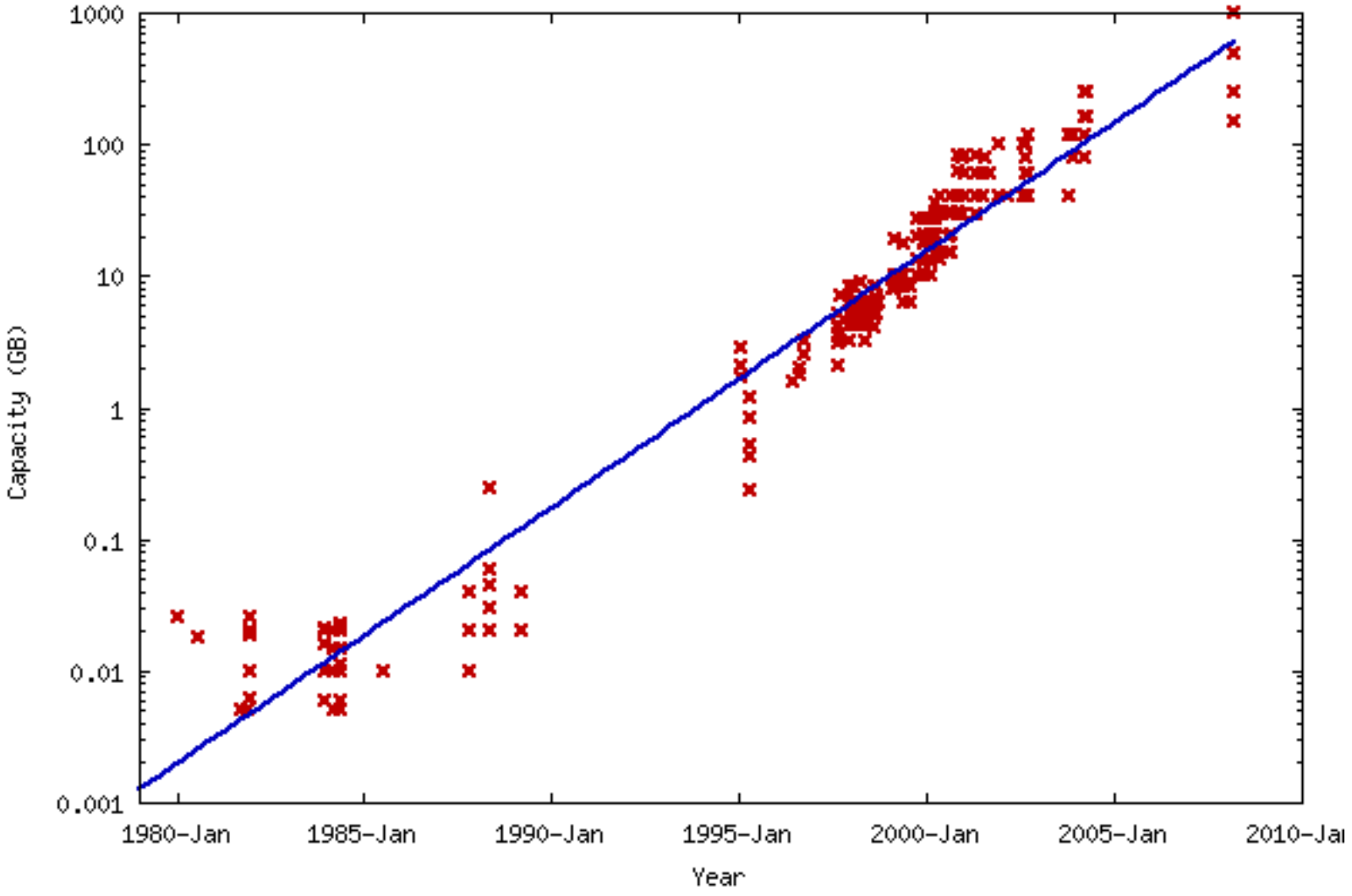
Software barrier

- A sometimes misunderstood point is that exponentially improved hardware does not necessarily imply exponentially improved software performance to go with it.
- We are hitting the software barrier before we hit the physical barrier
- Right now we have **multicore and multi-thread chips** inside computers, but if the software does not take advantage of them, you gain **no benefit**."
- Change programming paradigms:
  - Parallel and multi thread coding
  - SIMD / MIMD …
  - ….

SOME REFS:

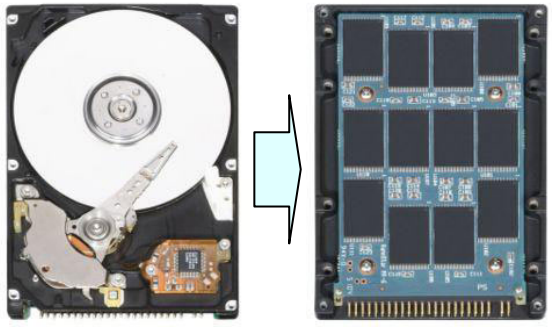
- http://www-unix.mcs.anl.gov/dbpp/
- http://www.mhpcc.edu/training/workshop/parallel_intro/MAIN.html ;
- http://www.fftw.org/parallel/parallel-fftw.html

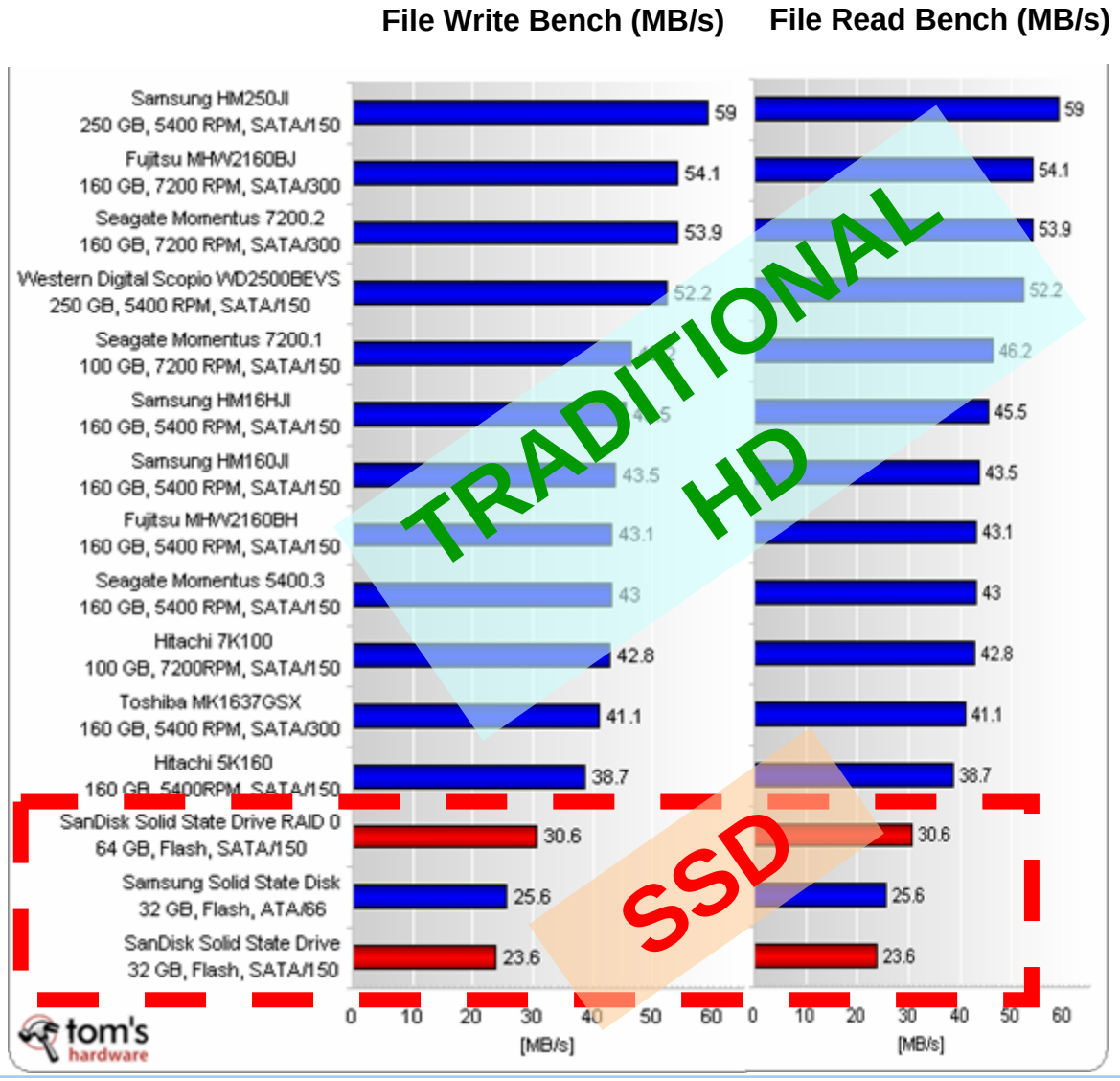
Storage Capacity evolution
Moore's Law..



Storage next step

A realistic future for the storage technology is realized by Solid State Drive (SSD)





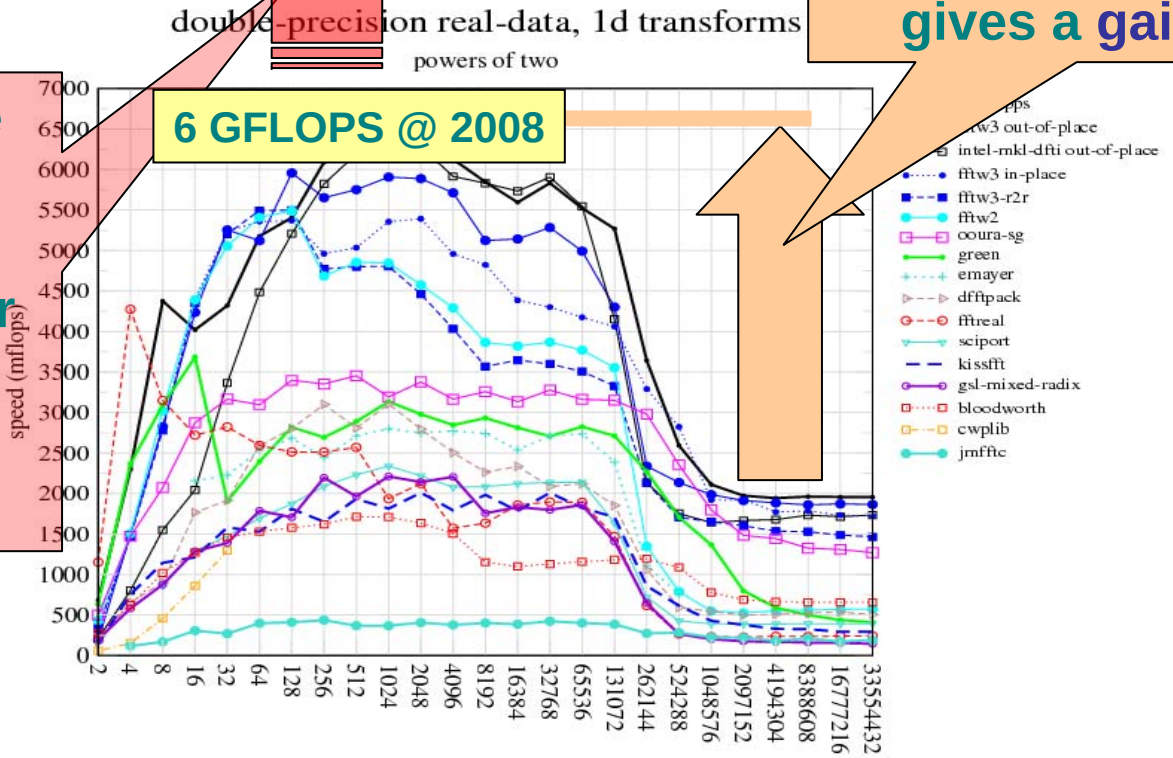
A possible 2020 scenario… the FFT

600 GFLOPS @ 2020 !!

Hyp1: In 2020 Memory IO gap could be compensated. This gives a gain factor=3

Hyp2: In 2020 the processor computational forecast power gives a gain factor = 100

6 GFLOPS @ 2008



A possible 2020 scenario… the FFT and CB detection case

- If we consider the matched filtering algorithm and the standard detection of a CB signal of:
  - 250s length (512 of buffer) @ 4kHz
- We need now (roughly):
  - 220ms for a single signal
  - And considering a bank of 4000 templates we need a total time of 880 s = 15 hours
- In 2020 we can try to forecast using the 300 gain factor:
  - 0.008 ms for a single signal
  - And considering a bank of 4000 templates we need a total time of **3 s !!**

The Virgo farm in 2020..

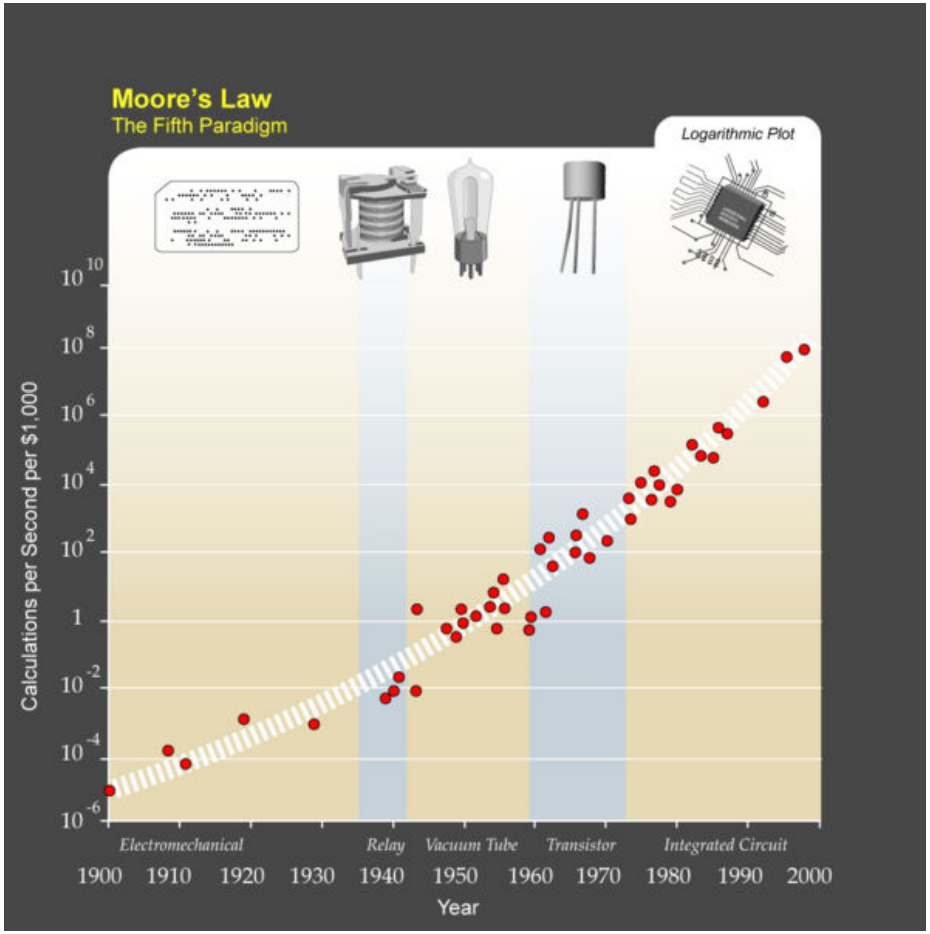
- If we continue to figure out the computing power available in 2020 and to provide a size of problem, we can use the Virgo farm as example.
- The actual Virgo farm is:
  - composed by more than 250 cores and 100 nodes.
  - computing power of about 400 GFLOPS
- In 2020 maybe we will have probably more than **500 GFLOPS on a single machine!**

Next decade processors technology

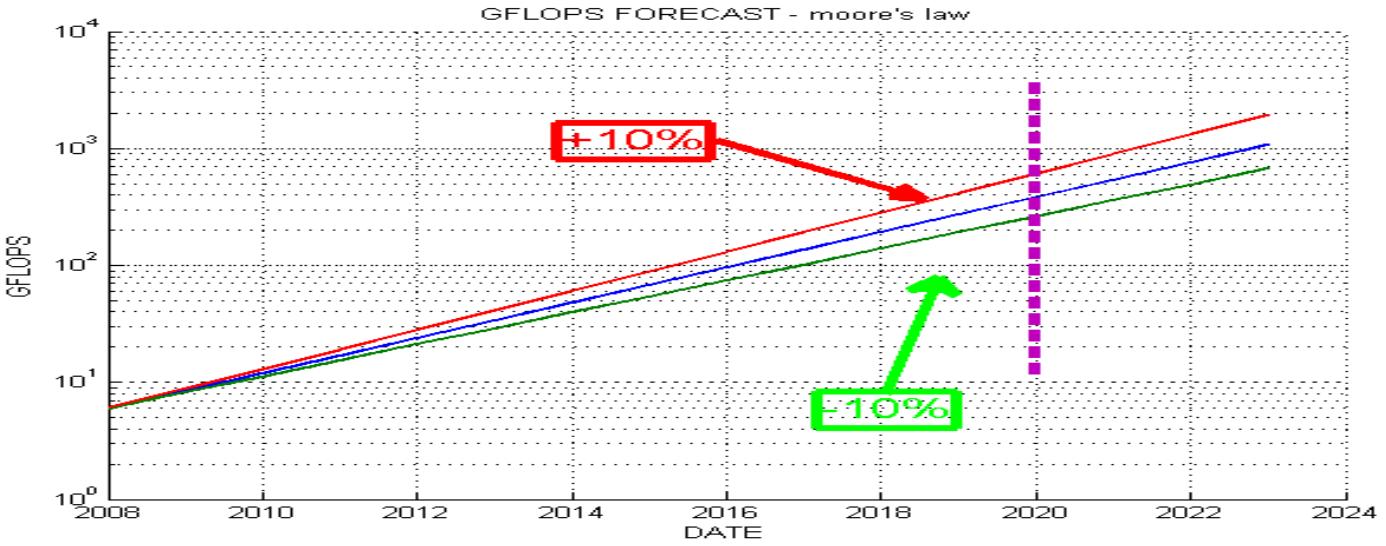
- **Chip-Level Multiprocessing (CMP)**
  - increasing parallelism for increased performance
- **Special Purpose Hardware**
  - Embed important functions relegated to software and specialized chips inside the microprocessor itself.
- **Large Memory Subsystems**
  - Memory access is a main bottleneck. In order to keep many high-performing cores, it is important to have a large quantity of memory on-chip and close to the cores.
- **Microkernel**
  - Microprocessors will need a sizable integrated intelligence, in order to coordinate all this complexity: assigning tasks to cores, powering up and powering down cores as needed
- **Virtualization**
  - Future microprocessors will need several levels of virtualization. virtualization is needed to hide the complexity of the hardware from the overlying software. The OS, kernel and software should not have to deal with the intricacies of many cores, specialized execution hardware, multiple caches, reconfiguration and so on.

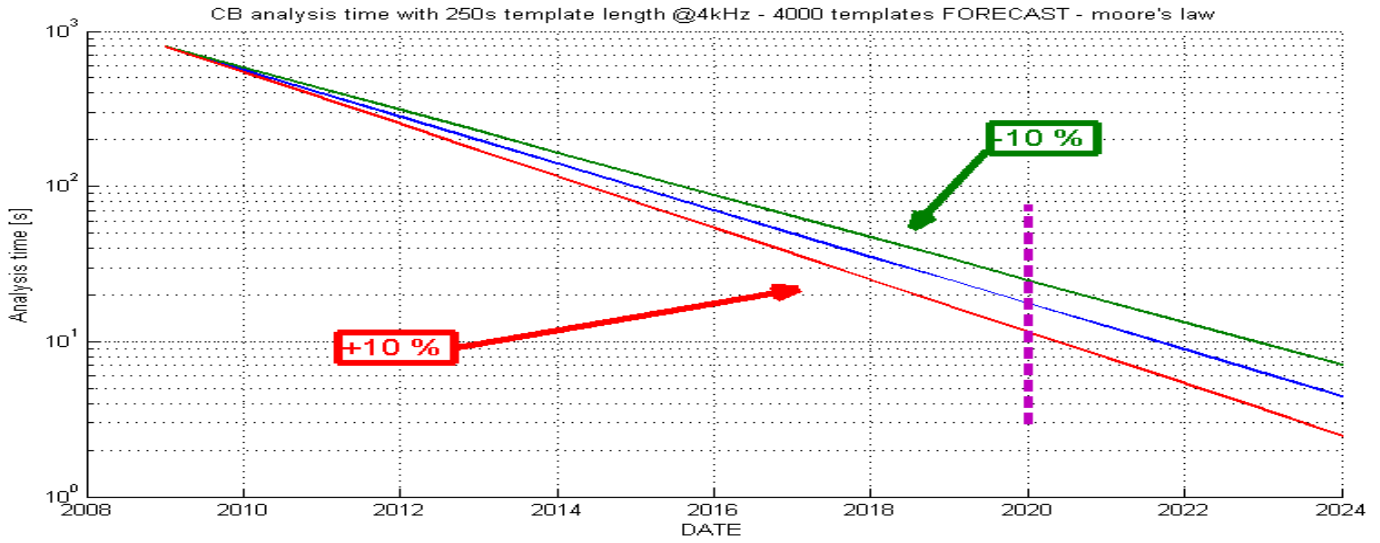
Conclusions…

- It is sure that around 2020 will take place some important technological changes. It is still not clear which will be the next technological step.
- A important effort in the scientific community is the CERN Computing Colloquia, that studies future trends in computing and information technology that are of interest to the physicist. http://cerncourier.com/cws/article/cnl/36011 (Thanks to Bruce A. for the infos about)
- Remember that now software is the main performance limitation. In the next years a programming paradigm change will be required.
- The future computing infrastructure will permit probably a real real-time/in-time GW observation needed by a GW telescope such as ET.



Extra


GFLOPS FORECAST - moore's law
GFLOPS
DATE
+10%
-10%
10^0
10^1
10^2
10^3
10^4
2008
2010
2012
2014
2016
2018
2020
2022
2024


CB analysis time with 250s template length @4kHz - 4000 templates FORECAST - moore's law
Analysis time [s]
DATE
-10 %
+10 %
10^0
10^1
10^2
10^3
2008
2010
2012
2014
2016
2018
2020
2022
2024