

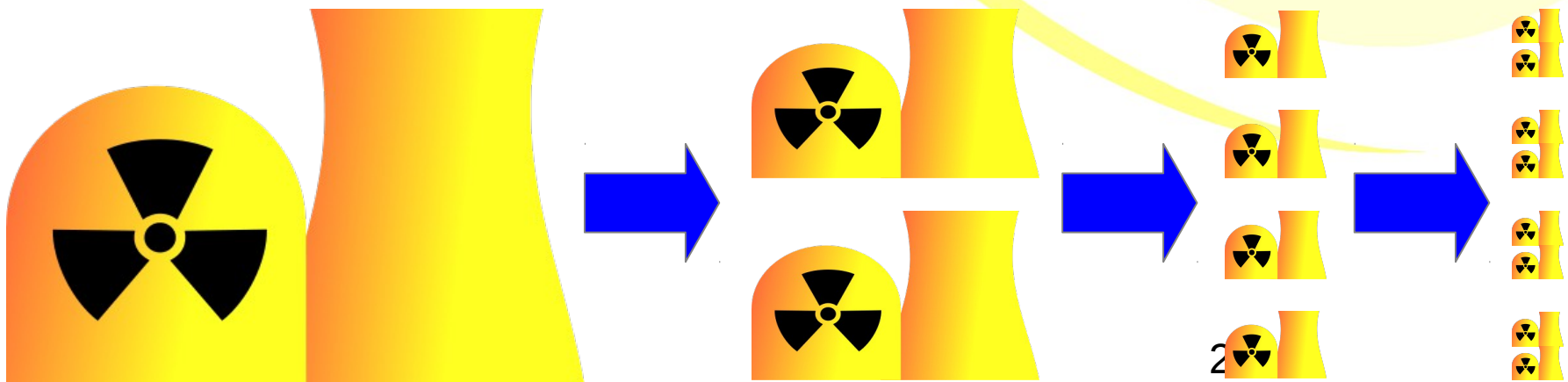
# Energy-driven algorithms *... or energy-aware computing...*

**Vincent Keller**



# Facts (2005)

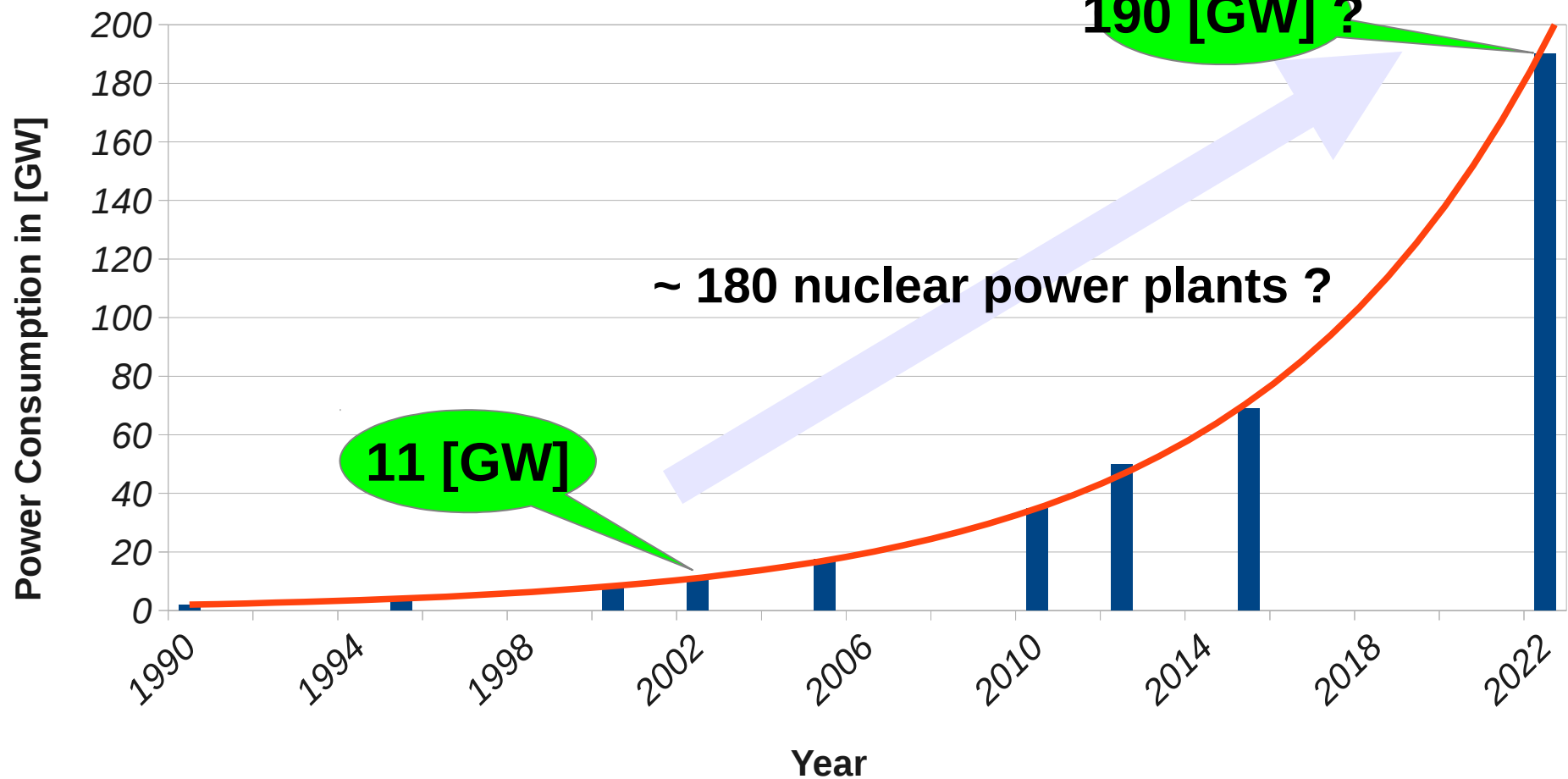
- Top 5 machines in the TOP500: 1-10 [MW]
- Power consumption for Data centers worldwide: 17,4 [GW]
- 1 mid-size nuclear power plant: 1 [GW]
- Power consumption needed by Data Centers doubles every 5 years



# (Pessimistic) Estimation 2002 – 2022

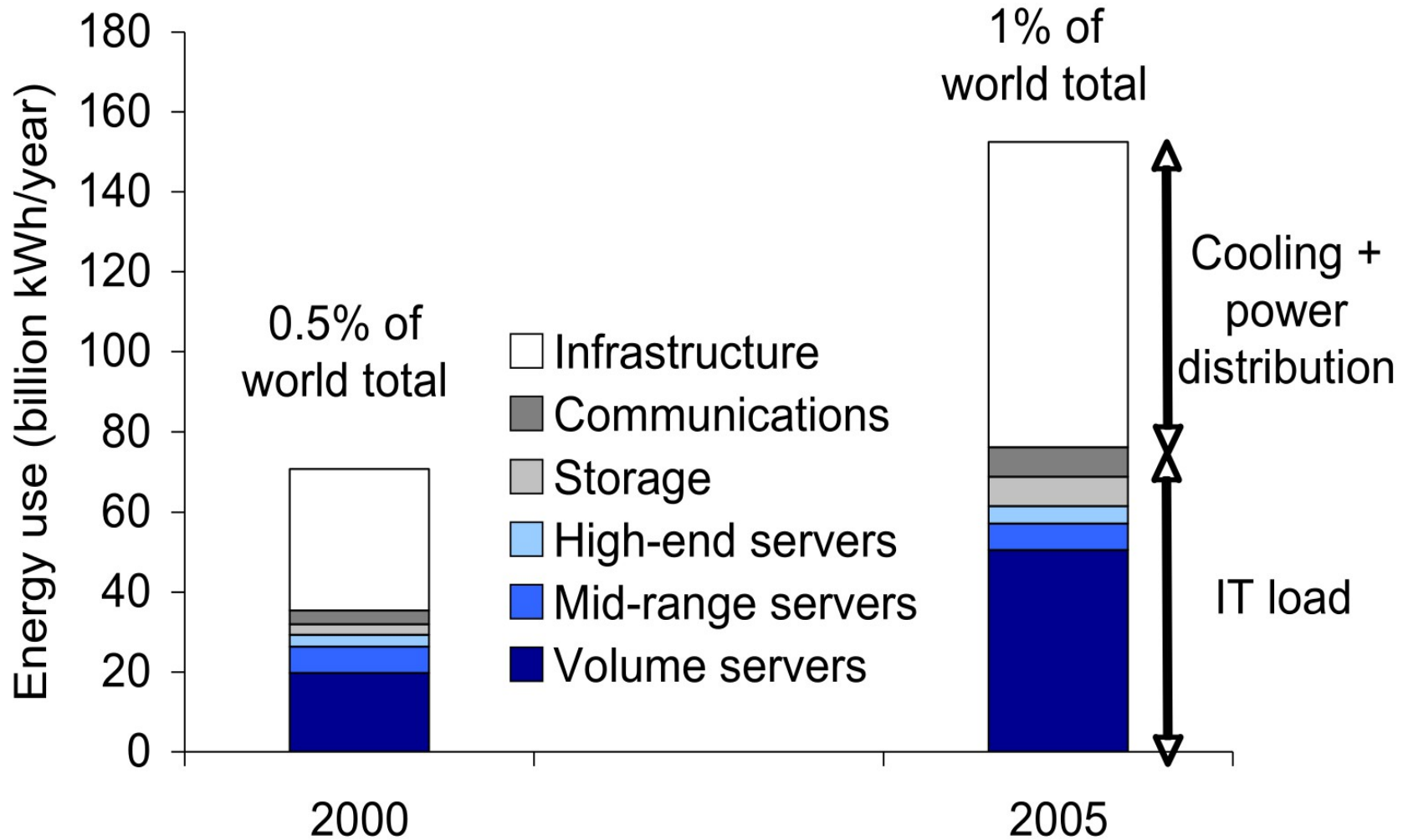
Power Consumption in Data Centers Worldwide

*(estimation based on Koomey, 2012)*



# Energy in Data Centers

## Where is it used ?



(Jonathan G. Koomey, 2012)



**Goal/Dream/Ads: Exascale ( $10^{18}$ ) in 2020**

**How ? (and why ?)**

# What vendors say

Home | TOP500 Supercomputing Sites - Iceweasel

File Edit View History Bookmarks Tools Help

Home | TOP500 Supercom... top500.org

Google

**TOP 500**  
SUPERCOMPUTER SITES

**SUPERMICR** 2012 NAB SHOW April 14-19<sup>th</sup> Booth # N2533  
**GPU SuperComputing**  
Intel Xeon ES-2600 HPC Systems Optimized for Scientific, Engineering and Computational Finance Applications  
Learn More >

PROJECT | **LISTS** | STATISTICS | RESOURCES | NEWS

CONTACT | SUBMISSIONS | LINKS | HOME

### TOP 10 Systems - 11/2011

1	K computer, SPARC64 Villix 2.0GHz, Tofu interconnect
2	NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050
3	Cray XT5-HE Opteron 6-core 2.6 GHz
4	Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU
5	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows
6	Cray XE6, Opteron 6136 8C 2.40GHz, Custom
7	SGI Altix ICE 8200EX/8400EX, Xeon HT QC 3.0/Xeon 5570/5670 2.93 GHz, Infiniband
8	Cray XE6, Opteron 6172 12C 2.10GHz, Custom
9	Bull bulk super-node S6010/S6030
10	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband

### Japan's K Computer Tops 10 Petaflop/s to Stay Atop

BERKELEY, Calif.; KNOXVILLE, Tenn.; and MANNHEIM, Germany (Nov. 14, 2011)—Japan's "K Computer" maintained the top spot on the TOP500 List of the world's most powerful supercomputers, a build-out that makes it four times as powerful as its nearest rival at the RIKEN Advanced Institute for Computational Science (AICS) in Japan, the K Computer it achieved an impressive 10.51 Petaflop/s on the Linpack benchmark using 705,024 SPARC64 processing cores.

» Read more

### About Ivy Bridges and Tegra Hitches

Tue, 2012-04-10 03:13 | whispers

Intel causes confusion with the Ivy Bridge's launch date, restructuring attempt and Nvidia launches rumors about the Tegra 3 (amongst other things) because of Apple.

» Read more

### About Armadas and Sinking Chips

Tue, 2012-03-27 08:50 | whispers

Until now, things haven't been going well for Intel in the TV market. A new media processor is supposed to help it gain a foothold with settop boxes, and concerning the television business, Intel has further grandiose plans. AMD could minimally increase its market share in the past year, but suffered severe losses in the server sector. The new Opteron 3200 family is supposed to turn the tides.

» Read more

### About Old Annoyances and New Challenges

Mon, 2012-03-12 13:50 | whispers

### ISC Think Tank Series

Sponsored by HPCwire

**The TOP500 - 20 Years Later**

Presented at ISC'12

HPC WITC

### ISC events

cloud computing

**ISC Cloud '12**

September 24 - 25, 2012  
Mannheim, Germany

Addressing Critical HPC Workloads

ATPR intel Xeon

**We are here at the first place !**

### League of TOP 1

**K computer**  
The K computer - named for the

Cloudy v...  
MiserWare Offers Free Power Measuring Utility  
Cray Brings In New CTO to Drive Supercomputing, Big Data Integration  
GPU Systems Releases GPU Compute Platform



# What vendors say (AI Gore version)

The screenshot shows a web browser window displaying the Green500 website. The browser's address bar shows 'www.green500.org'. The website header includes the 'THE GREEN 500' logo, a navigation menu (HOME, ABOUT, GREEN LISTS, NEWS, FAQ, CONTACT), and a 'STATISTICS SEARCH' section with dropdown menus for '2011/11', 'GROUPINGS', and 'VIEWS', along with a 'DISPLAY' button. A large banner image features a green field, a blue sky with clouds, and a server rack, with the text 'Ranking the World's Most ENERGY-EFFICIENT SUPERCOMPUTERS'. Below the banner, the main content area is titled 'Environmentally Responsible Supercomputing' and contains introductory text about the Green500's mission. A yellow speech bubble is overlaid on the page, containing the text: 'We "should" be here at the first place !'. At the bottom of the page, there is a green box with the text 'GREEN500 THE MOST POWERFUL SUPERCOMPUTERS RANKED BY ENERGY EFFICIENCY.' and a section titled 'The June 2012 Green500 list will be open for submissions shortly.' with a 'submissions' link. To the right, there is a section titled 'Run Rules and Submission Portal Opening for the November 2010 Green500 list' with a date of 'Oct 19, 2010' and a link to the 'Run Rules and Submission Portal Opening'.

**The vendors base all their forecasts (and R&D) on the performance of HPL (High Performance LINPACK) for the TOP500 and GREEN500 lists**

***HPL is a highly optimized application's kernel***

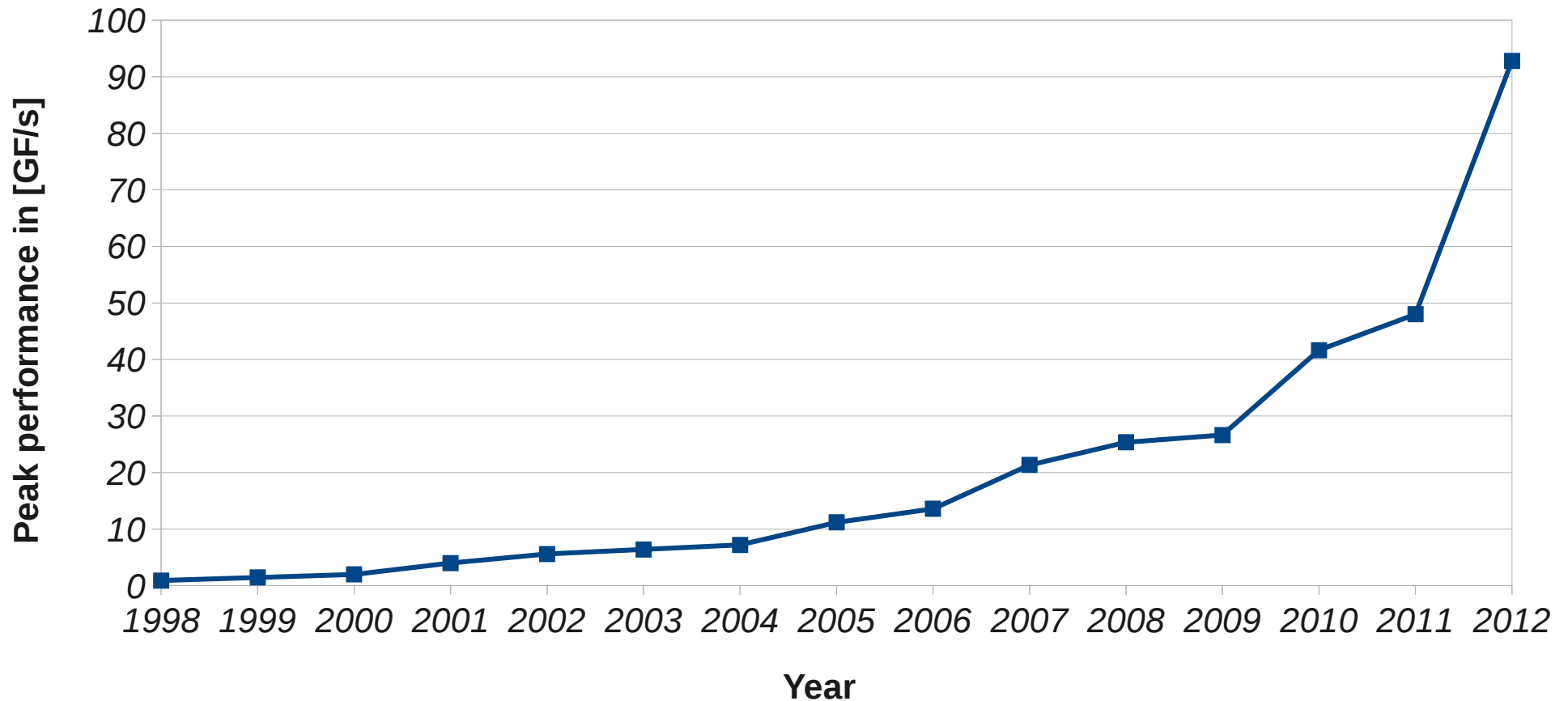
***HPL is purely CPU-bound: its performance ( $R_{max}$ ) is very close to the CPU peak performance***



# Processor evolution ( $R_{\max}$ )

## The Intel example

Peak performance of Intel processors  
*From 1998 (Pentium II/XEON) to 2012 (SandyBridge)*





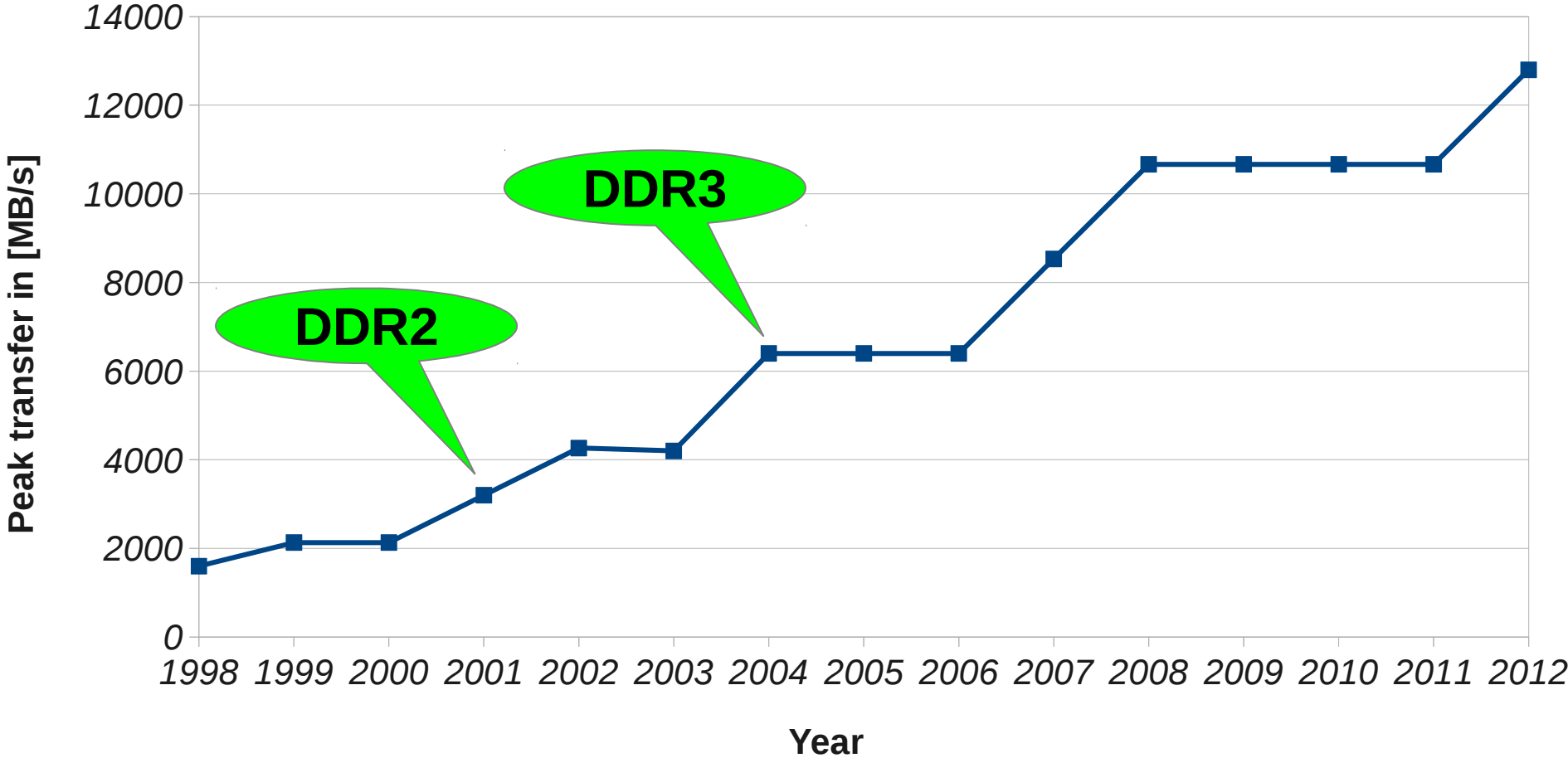
**OK, the CPU performance follows the Moore's law for HPL.**

**What about the memory ?**

# Memory performance evolution

## Memory peak transfer rate per link

*From DDR (1998) to DDR3 (2012)*





**Nice correlated data, isn't it ?**

**No ! CPU performance grows exponentially  
while memory linearly.**

**So ... are our applications CPU  
or memory access dominated ?**

$$Ax = b$$

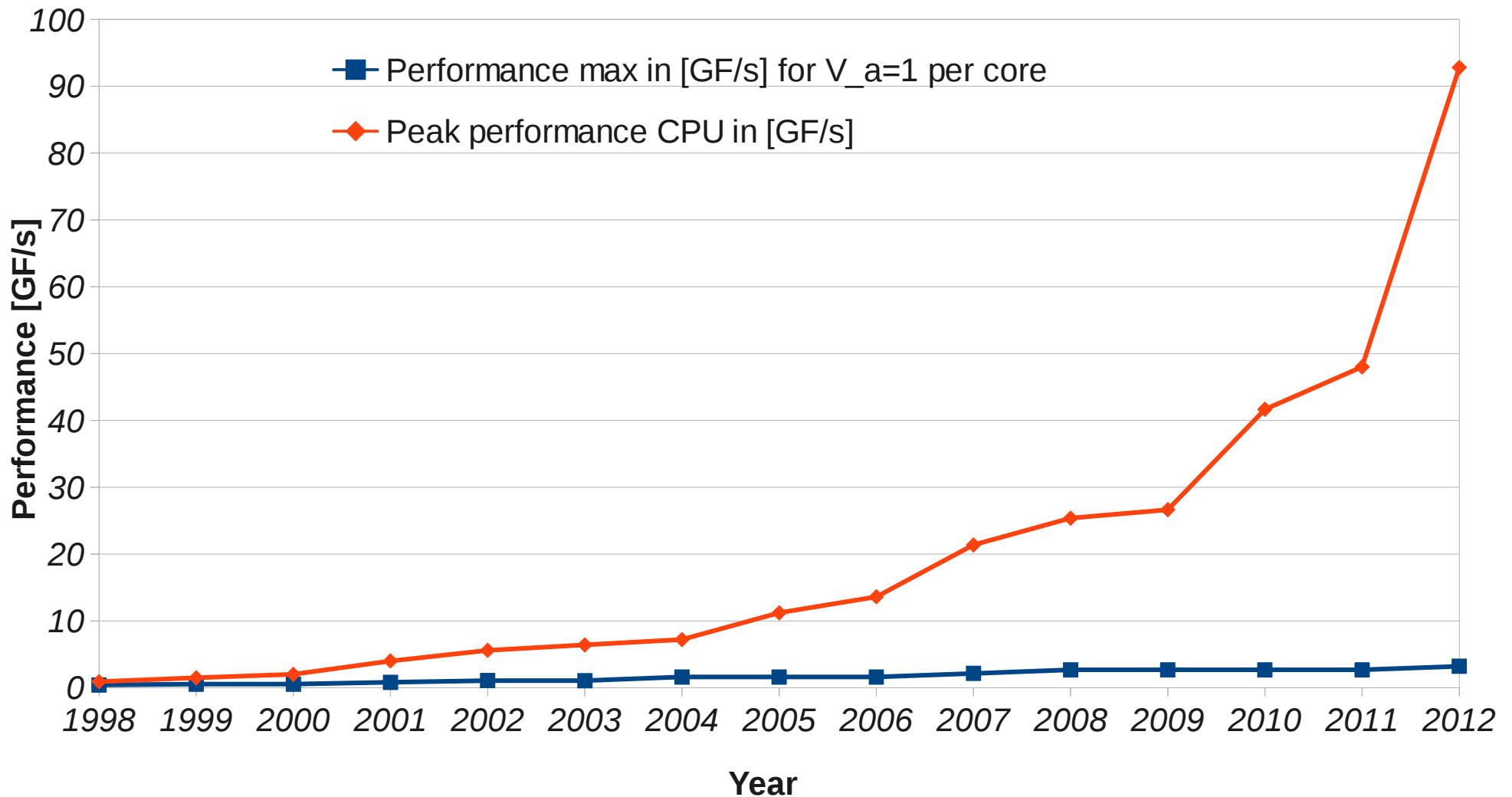
- Many simulation codes can be “summarized” to the resolution of  $Ax = b$ .  $A$  is (probably) sparse and  $\dim(x) \ll \text{large enough}$  »
- Let me define  $V_a$ , the “vectorization ratio”: the number of operations that can be performed by the CPU per memory access (LOAD or STORE)
- If  $V_a = 1$ , each operand must be loaded from memory.
- This is the case for  $Ax = b$



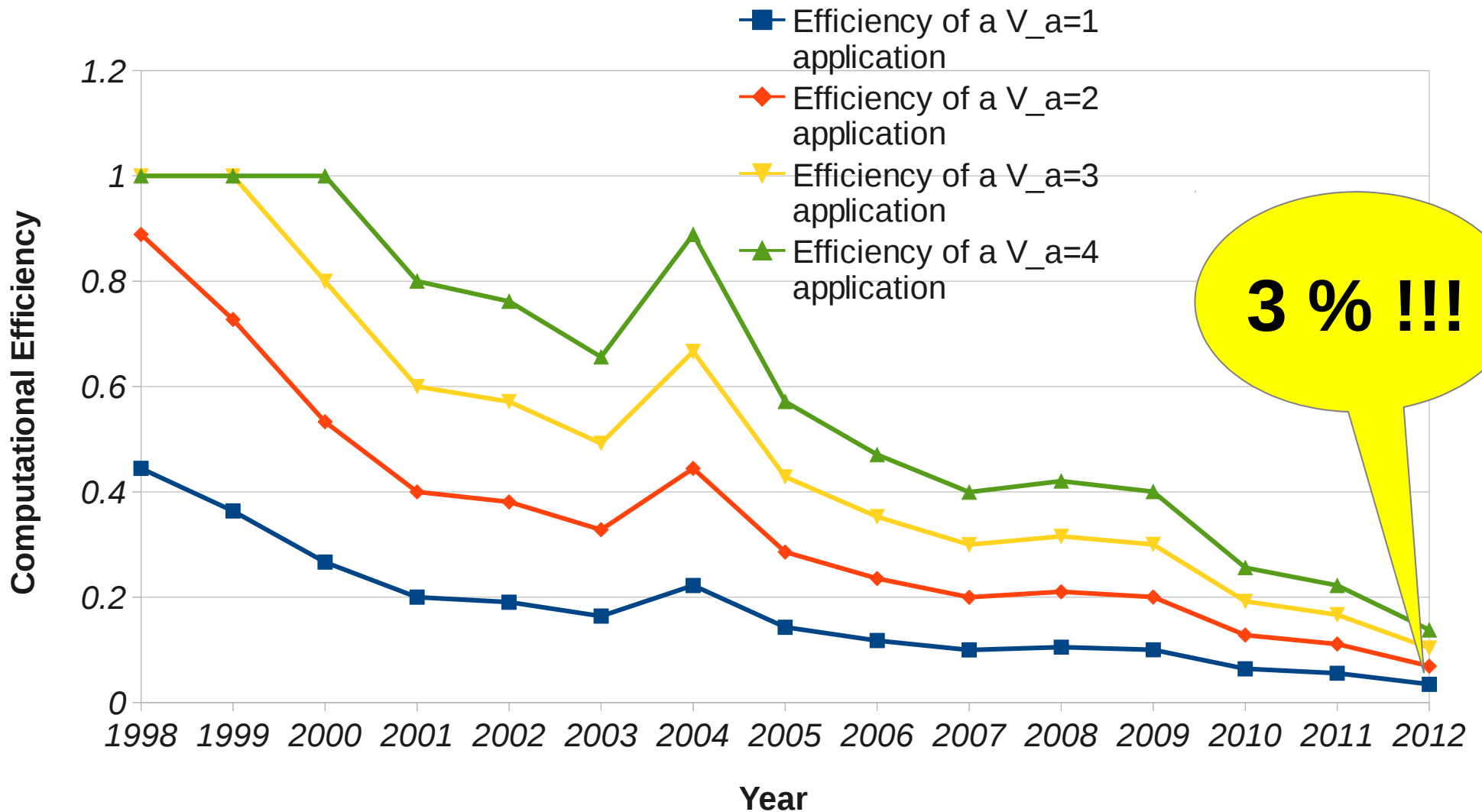
# Consequence

- If  $V_a = 1$  (or close), the application is purely memory bounded.
- Memory bandwidth and latency are the bottleneck
- CPU performance is not an issue

# Comparison HPL and a memory-bound application



# Computational efficiencies of a $V_a=\{1,2,3,4\}$ application



# Energy speaking...

- 3% of computational efficiency means:
  - 3% of the CPU is used for what it is built for
  - 97% of the CPU is not used (but the CPU does not stop)
- 97 % of the energy brought to the CPU is used to heat the air (or the water) around the data center

**And we want to go for Exascale ?**



**OK, that is for a “pure”  $V_a=1$  application.**

**What about REAL applications Energy footprint ?**



# Test case

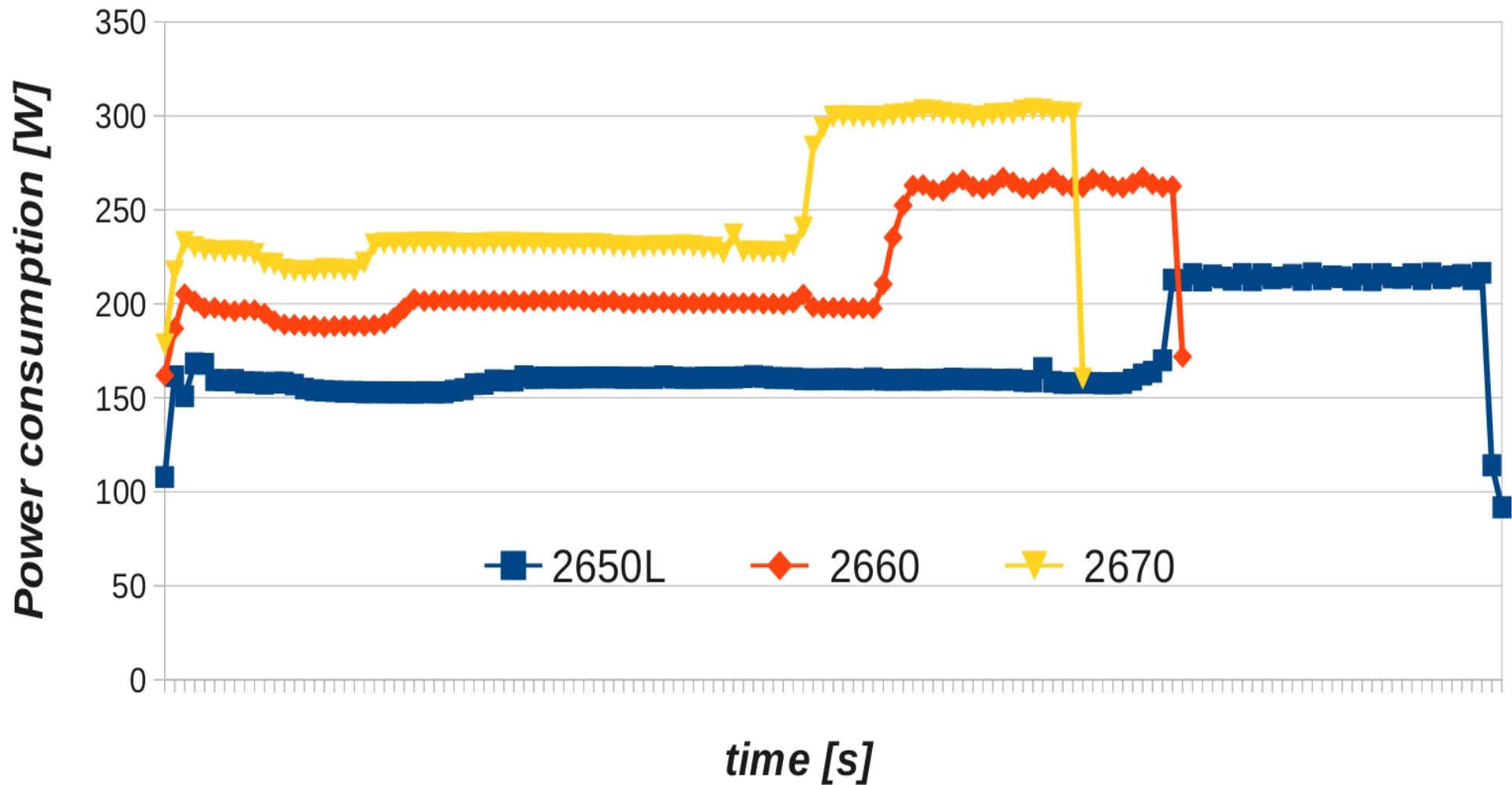
- Three applications
  - MiniFE (FEM)
  - CPMD (Car-Parinello)
  - GEAR (n-body)
- Three 2012 CPU models:
  - CPU1 is “low power”
  - CPU2 is “middle class”
  - CPU3 is “high performance”
- All three CPUs have the same instruction set and the same memory banks connected.



**Power consumption is measured during execution...**

# An example: SNL MiniFE

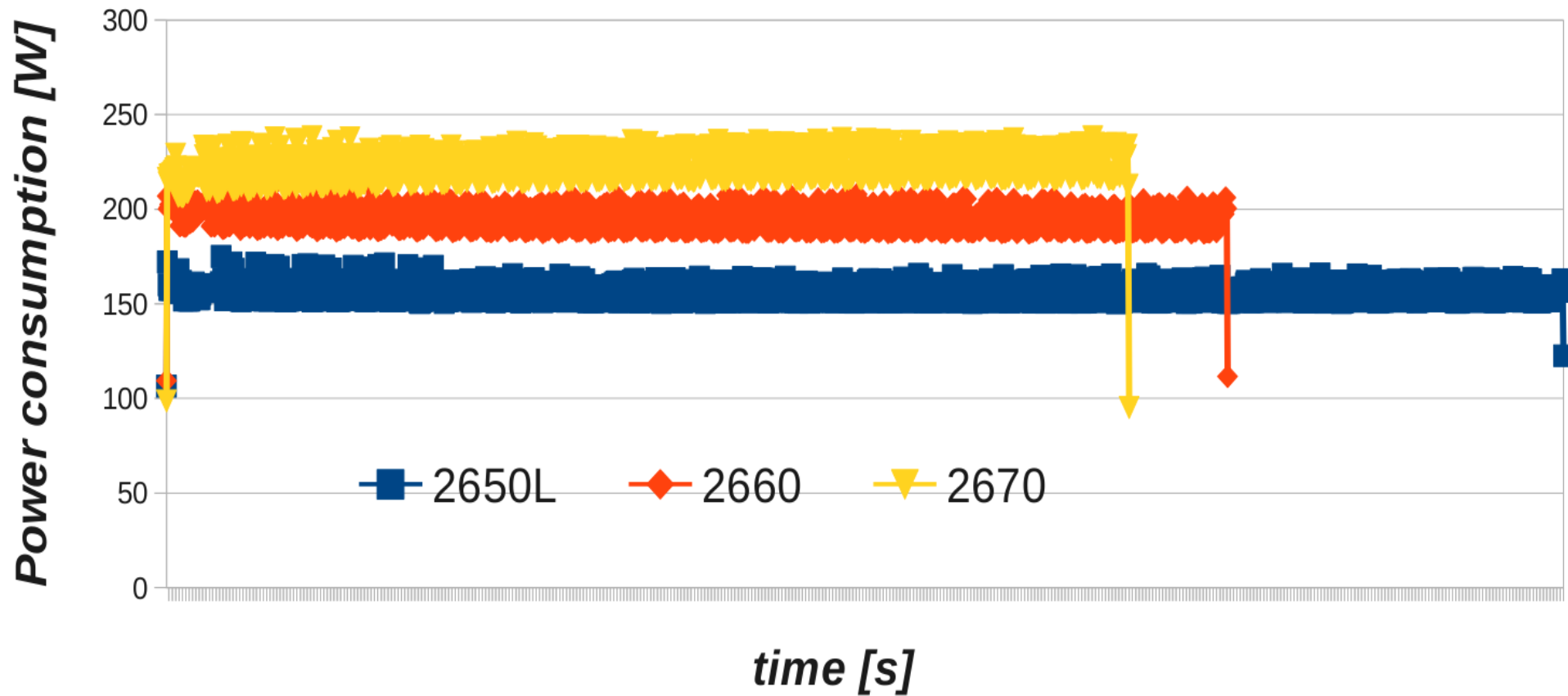
SNL MiniFE on 8 cores



MiniFE: “best approximation to an unstructured implicit finite element or finite volume application, but in 8000 lines or fewer.”

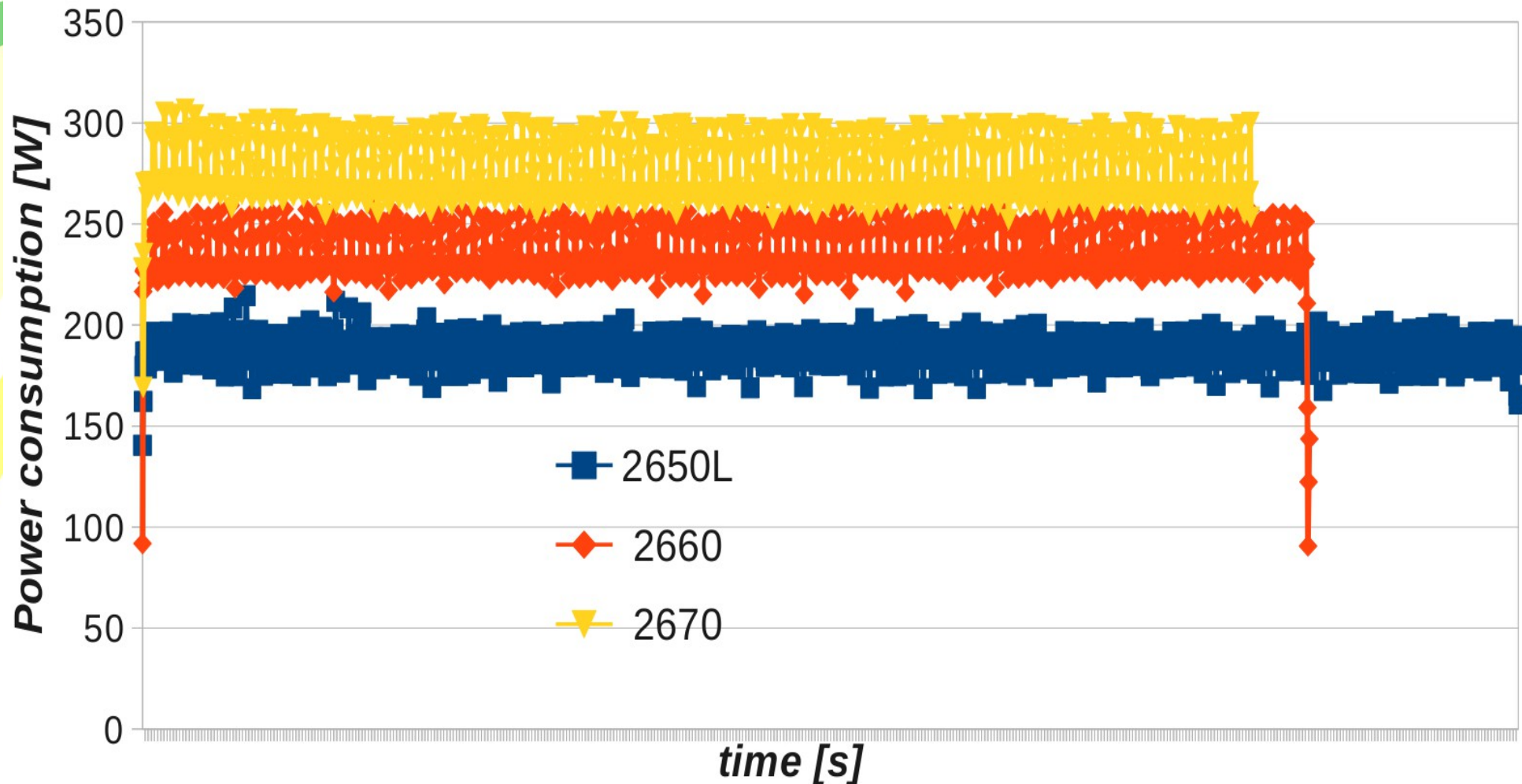
# Another example: n-body GEAR

GEAR on 8 cores



# A last example: Car-Parinello CPMD

CPMD on 8 cores





# Energy To Solution

- $E = P \cdot t$  ( $E = \int p(t) dt$ )
- Energy-driven computing → choice of the most efficient machine for a given algorithm
- TTS (Time-to-Solution) can differ. We strictly deal

with Energy

# Energy-to-solution (in KJ)

Application	Number of cores	CPU1	CPU2	CPU3
MiniFE	2 cores	54.01	<b>48.225</b>	48.823
	4 cores	29.571	<b>27.695</b>	28.672
	8 cores	23.072	<b>22.206</b>	23.087
	16 cores	<b>18.547</b>	19.425	21.761
GEAR	1 core	1223.269	1010.254	<b>982.307</b>
	2 cores	734.507	644.846	<b>636.732</b>
	4 cores	420.228	<b>382.701</b>	385.091
	8 cores	258.579	<b>246.42</b>	254.397
	16 cores	<b>224.916</b>	231.221	252.747
CPMD	4 cores	666.9	<b>658.952</b>	698.769
	8 cores	<b>407.248</b>	433.001	477.28
	16 cores	<b>350.794</b>	364.845	467.274



So HAL, what is your conclusion?

“I am putting myself to the fullest possible use,  
which is all I think that any conscious entity can  
ever hope to do. “

HELLO DAVE

# Conclusion

- Choose the best suited machine.
  - Ex: For a memory-bound application ( $V_a=1$ ), a **vector architecture** is more suited
- Change the algorithm if your problem can be solved differently
- Help CPU founders to produce suited chips with a balanced FSB/Core frequency: **CO-DESIGN**

# Conclusion

- Exascale in 2018 : probably possible.
- But (if we want to use it) :
  - Not with the current technology
  - Not without a tremendous improvement in the memory performance
  - Energy is not free. Energy-to-Solution metric will become more and more important with respect to Time-to-Solution
- Do we need that much power to make Science ?