

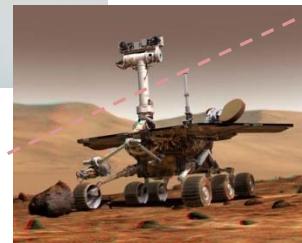
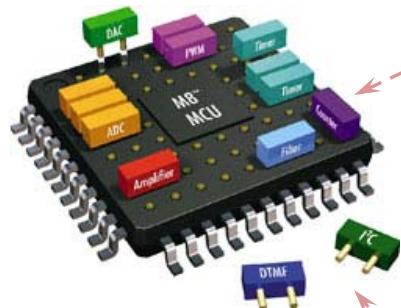
Modern application requirements



MOBILE



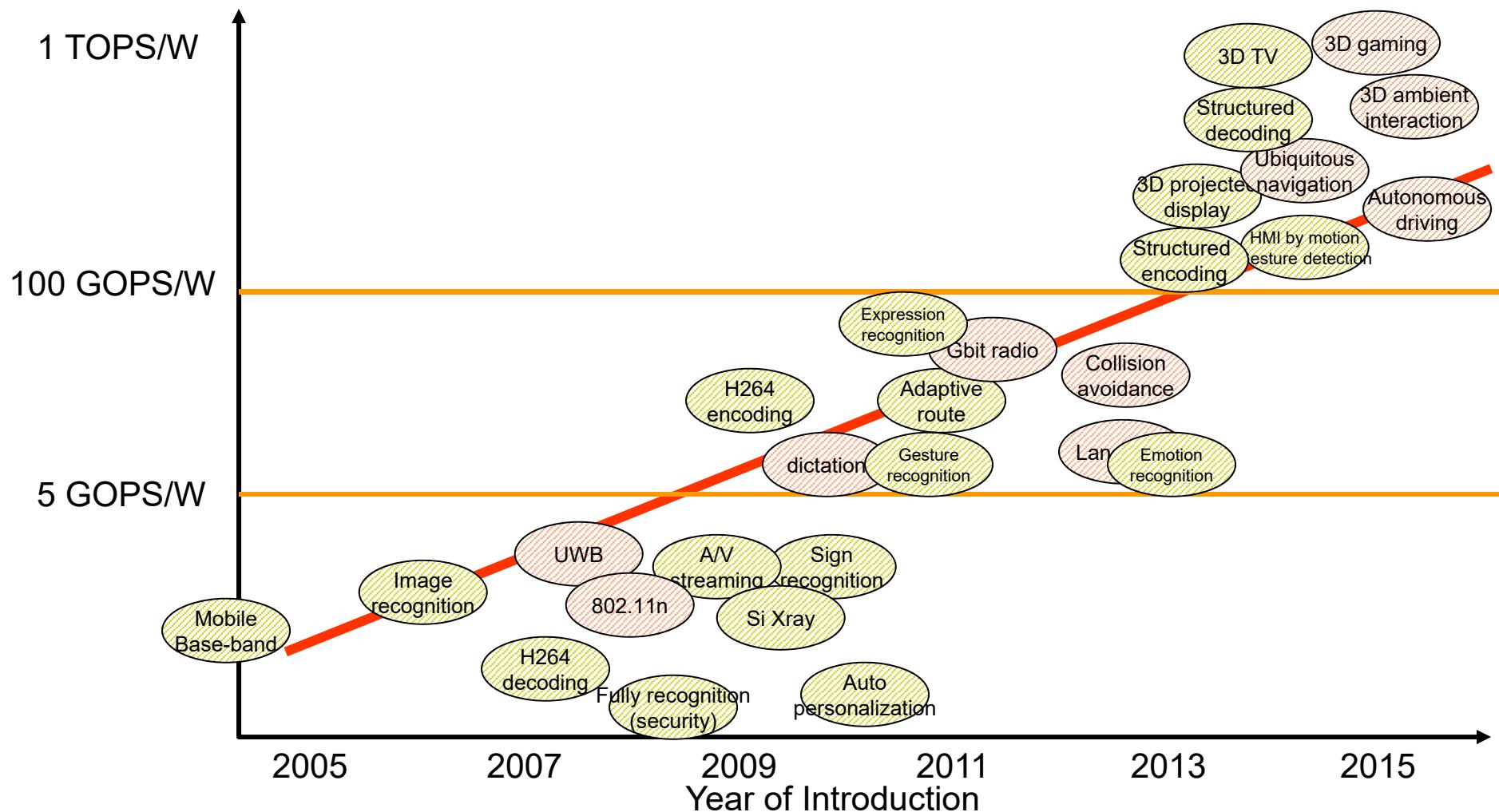
LIFE-CRITICAL



PROCESSING
POWER

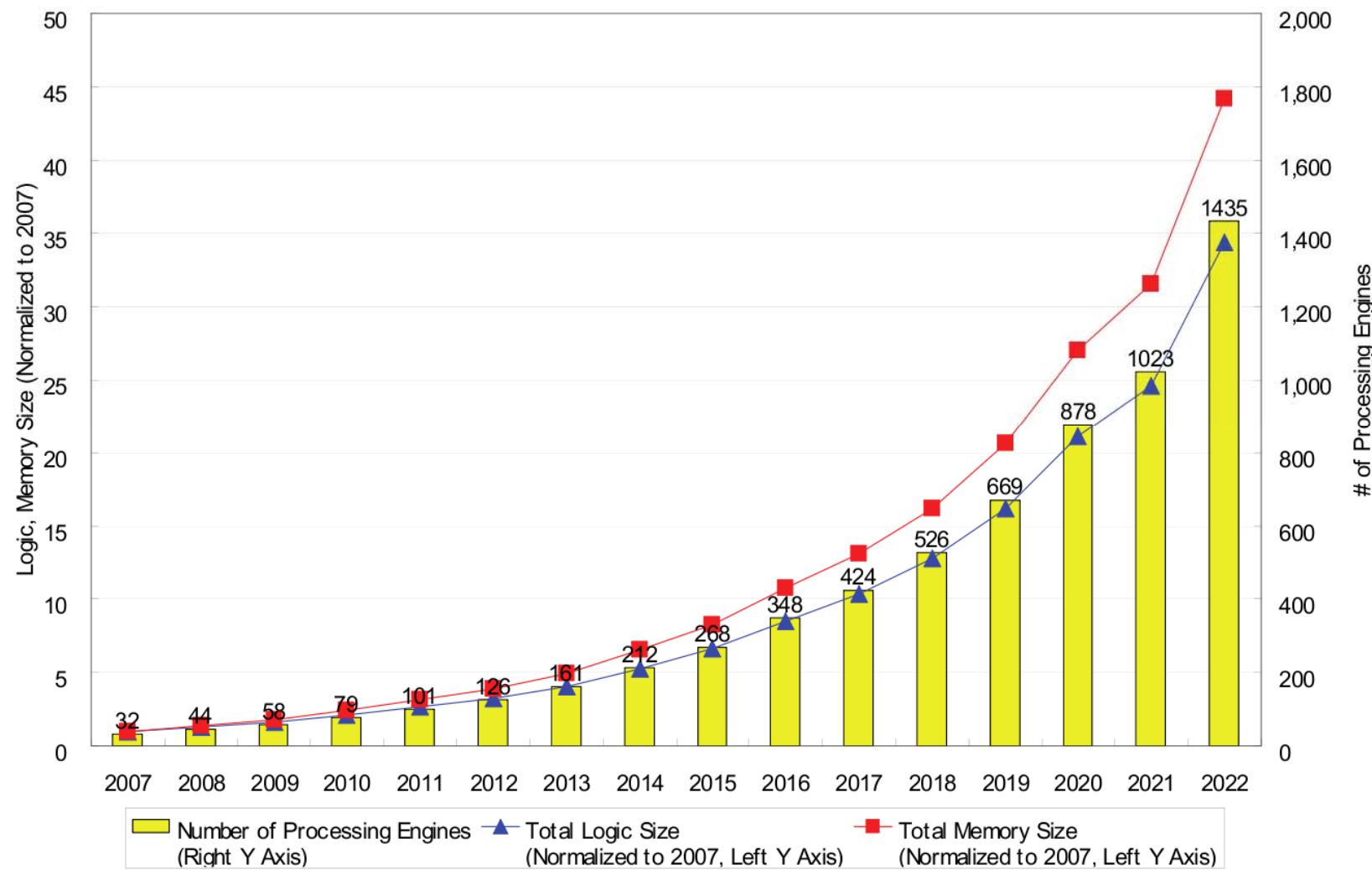


Market application rush

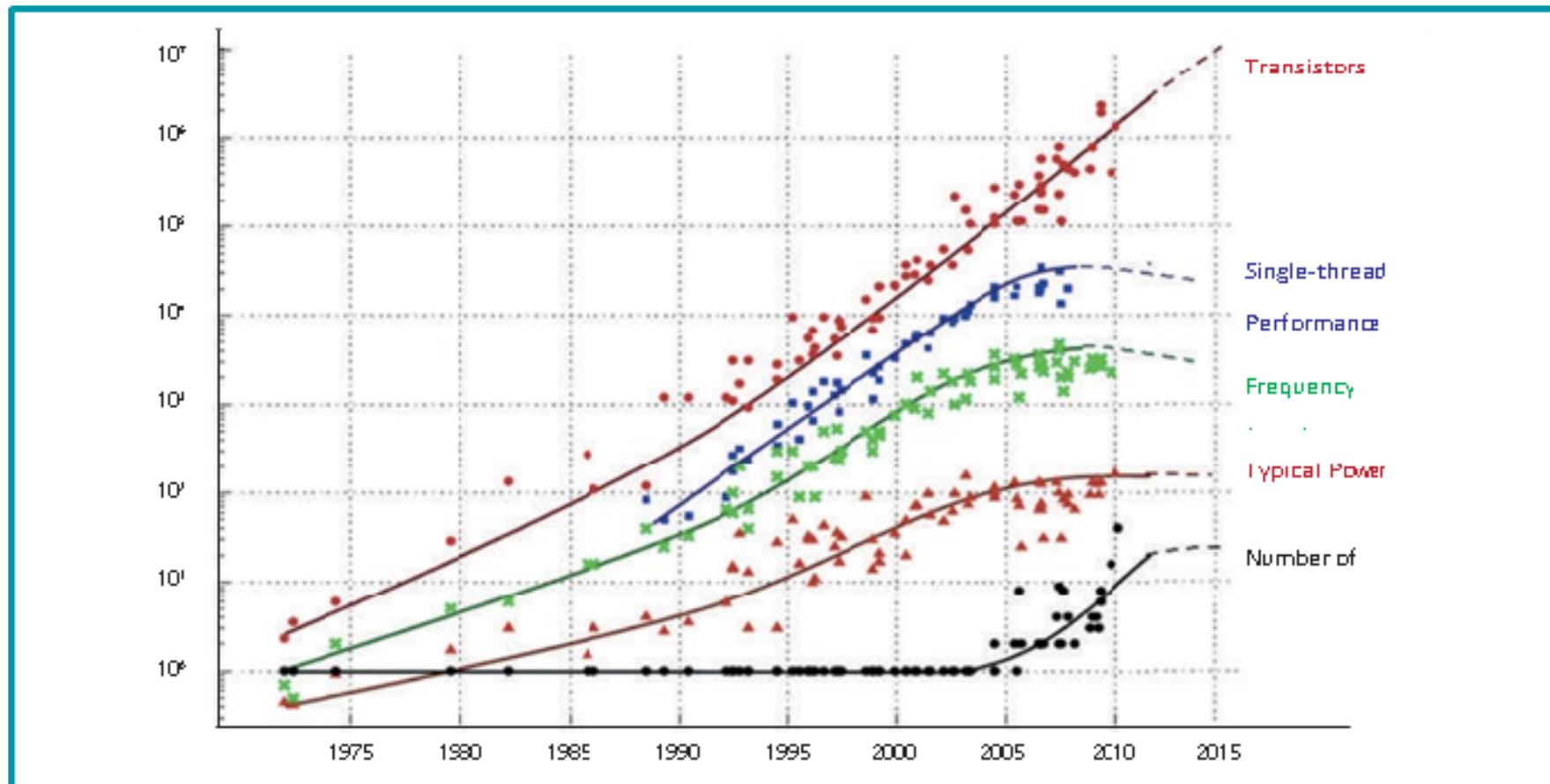


Increasing processing demand

- MPSoC architectures evolution
 - Many core, multi core, heterogeneous computing



Towards the dark silicon

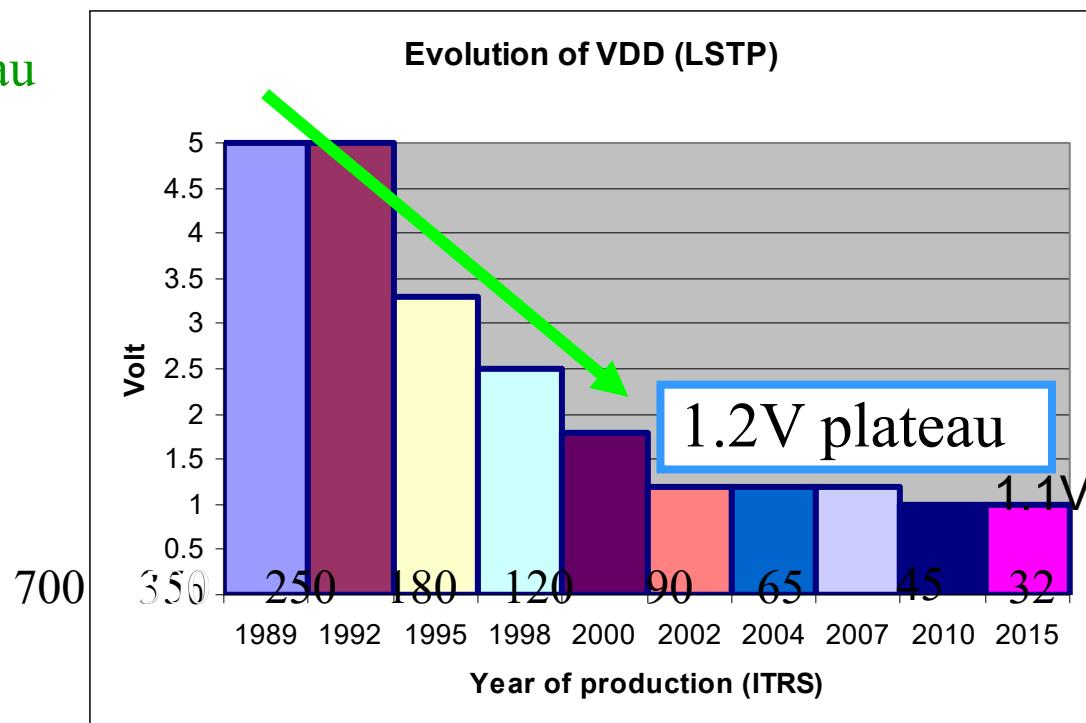


- Not-exploitable computing power due to limited power dissipation
- Part of the silicon area is ...»dark silicon»

VDD is no more scaling down

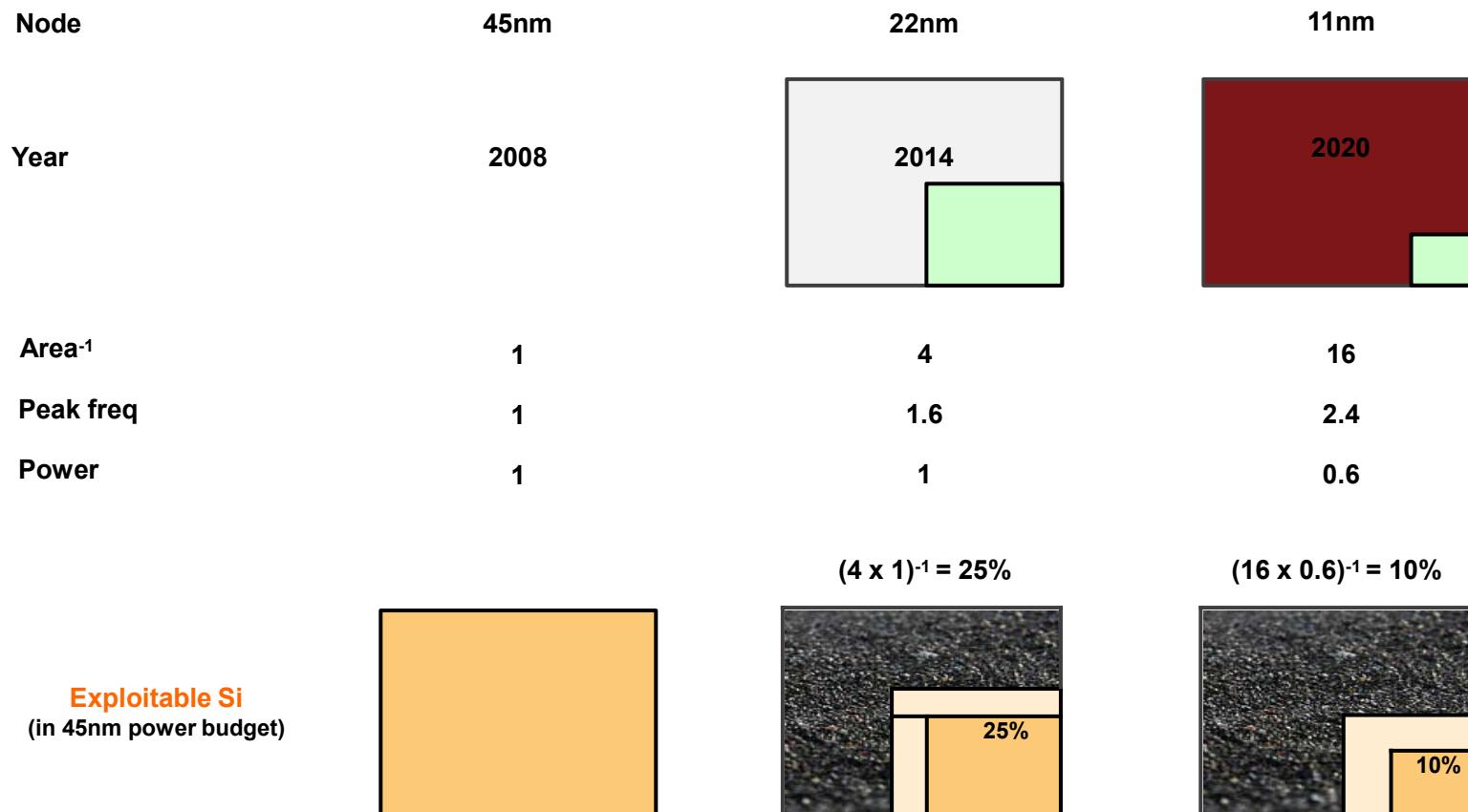
Regular decrease
5V to 1.2V (0.7x per node)

5V plateau



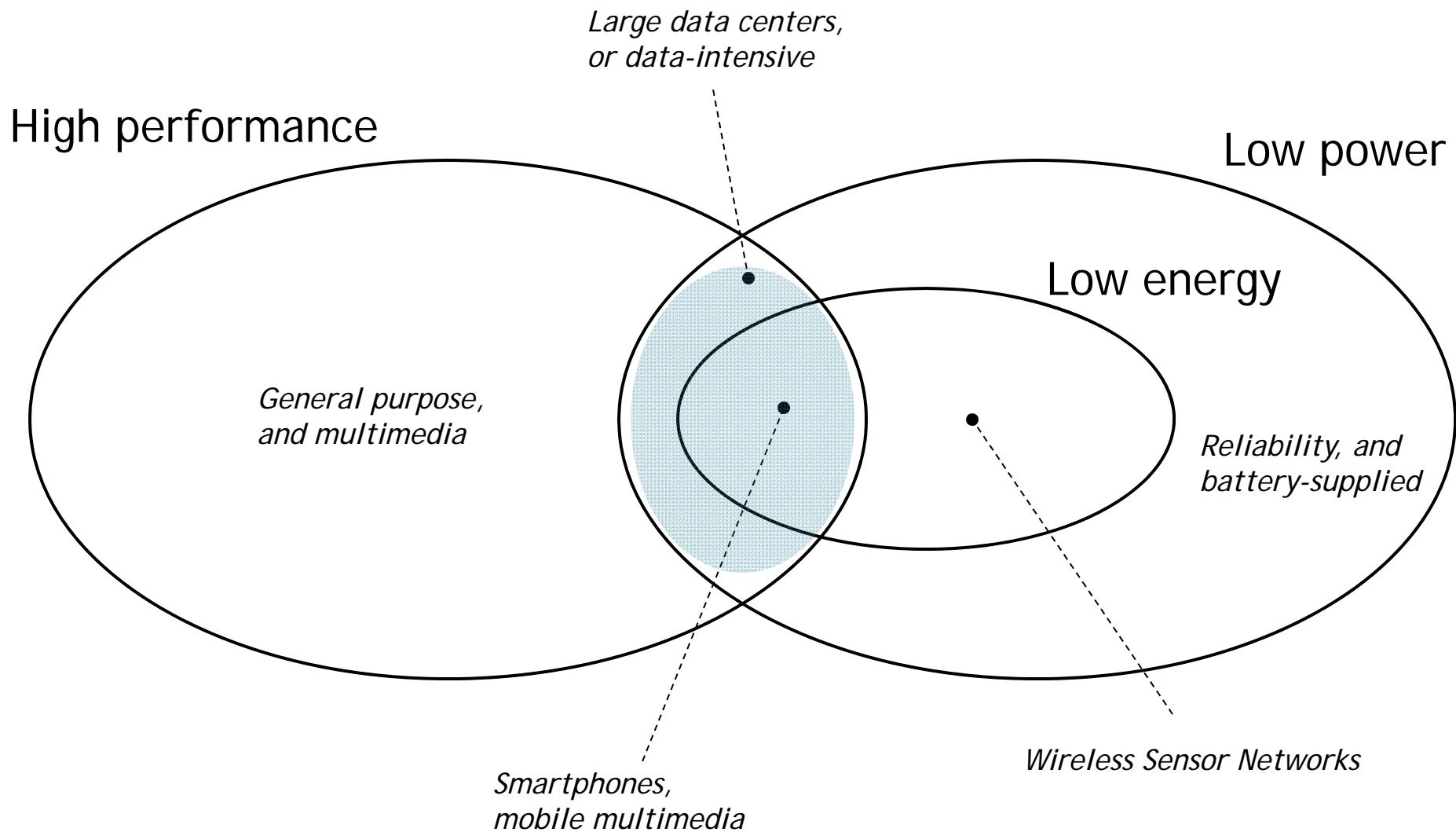
1V plateau?

Creation of dark silicon



Source: ITRS 2008

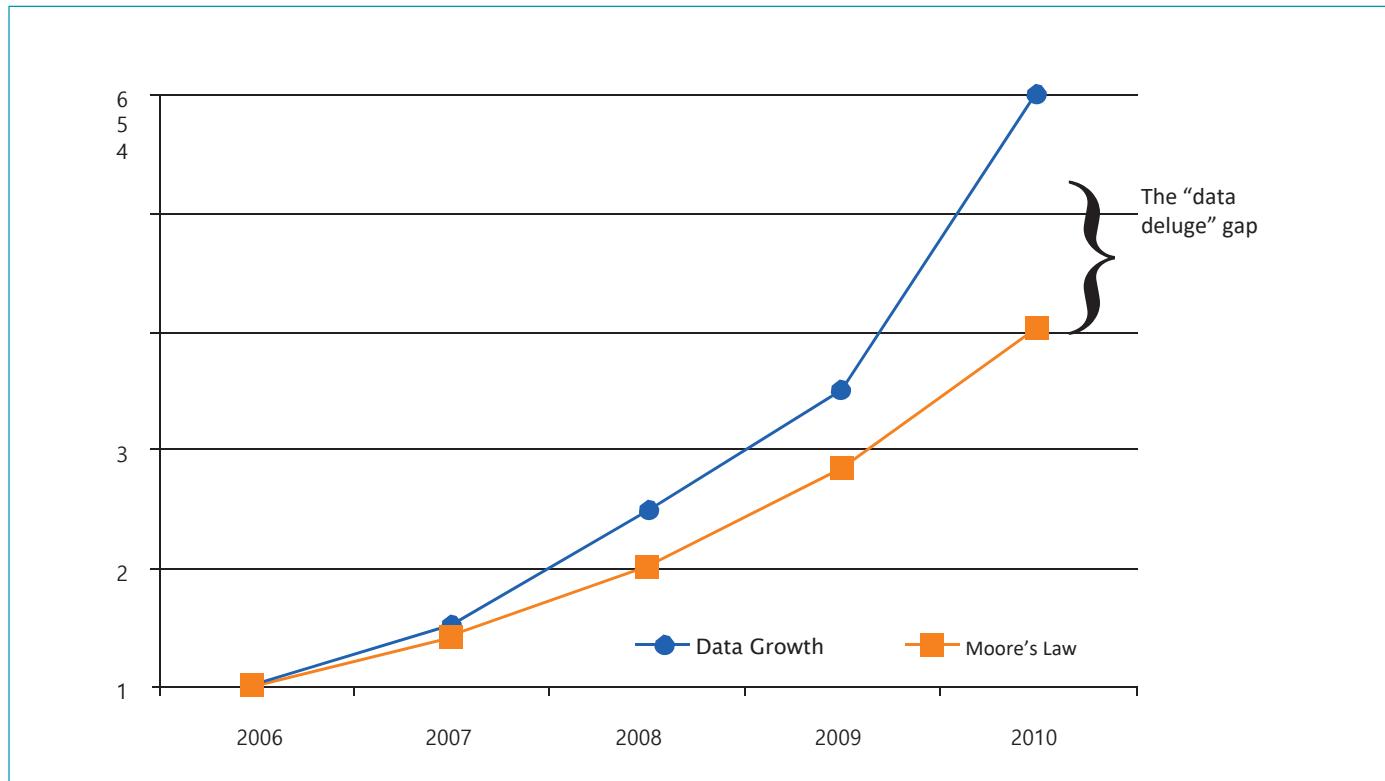
Application scenarios



What we can do

- We can have more transistors
- We just can't power them all at the same time
- We need to use these extra transistors in new ways
 - Multicores
 - Many-cores
 - Domain-specific processors
- It all points to heterogeneous processing
 - And aggressive power management
- Computing to be done in the most efficient place

Now you have the bicycle...



- Data growth vs. Moore's Law trends in the last 5 years
- Data “deluge” means that we are heading towards a world where we will have more data available than we can process

The Networks-on-Chip Reliability Wall

Small transistors = Big problems

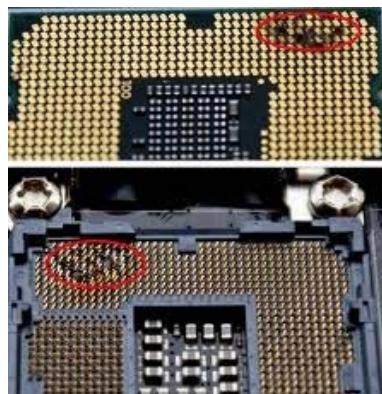
- Process Variation
- Physical Failure
- Aging mechanisms (NBTI)
 - (device performance decreases over years)



Small transistors = more packed transistors

(VLSI integration)

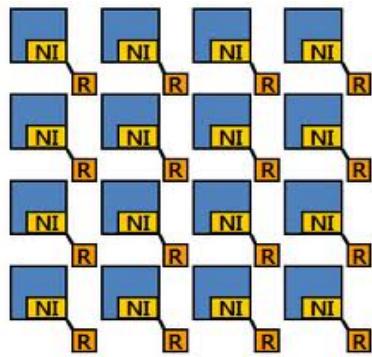
- Increased power density
- Thermal issues



“Reliability will be a barrier to future scaling”
Shekhar Borkar,
Intel Fellow

“Reliability will be a first class design constraint”
Chuck Moore,
AMD Senior Fellow

Network-on-Chip and Power-Performance



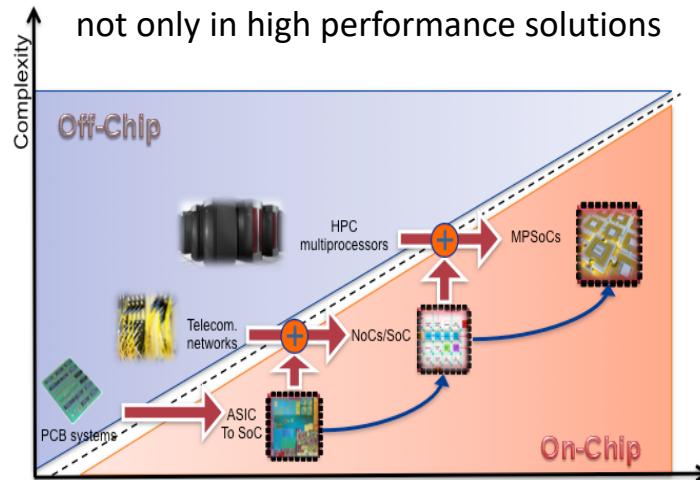
Network-on-Chip (NoC): multi-core flexible/scalable interconnect [1]

traditional communication subsystems cannot ensure adequate power/performance trade-off (buses, P2Ps, crossbars)

- *NoC power can be up to 30% of the total chip [3]*
- *NoC performance greatly influences the multicore [1,3]*

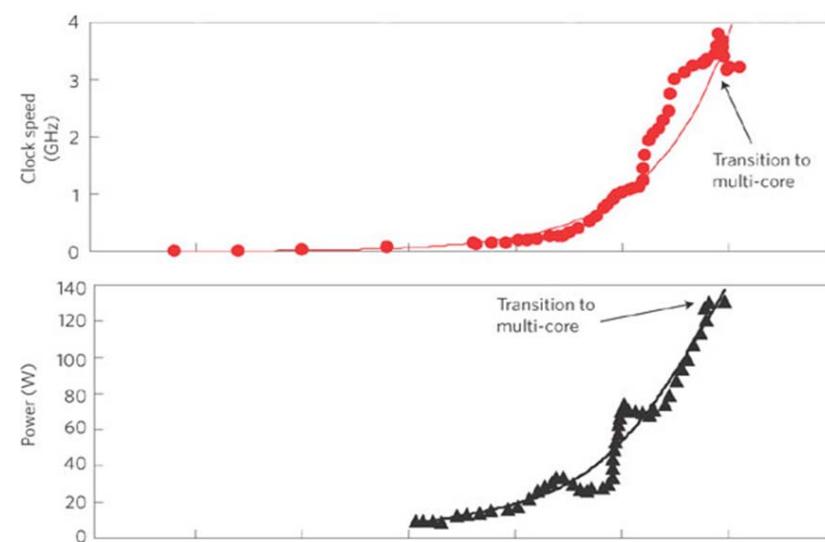
Single-core → Multi-core architectures

there is a need for even more performance
not only in high performance solutions



Power wall

performance are not free. Multi-core to optimize power performance trade-off [2]

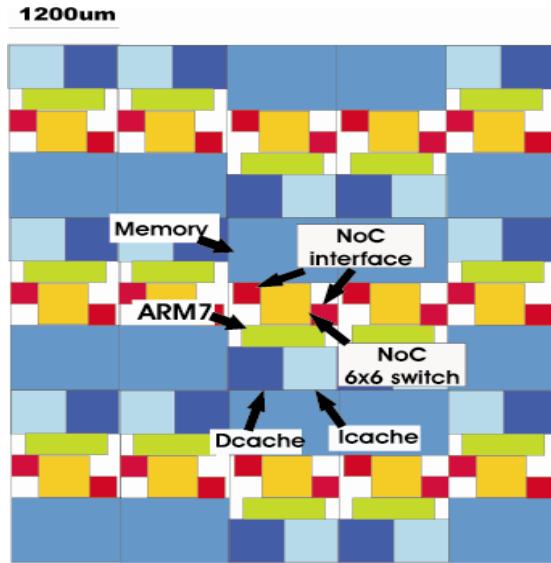


[1] G. de Micheli and L. Benini. Networks on chip: A new paradigm for systems on chip design. In DATE '02, page 418, Washington, DC, USA, 2002.

[2] A. Majumdar. "Helping chips to keep their cool". Nature Nanotechnology, April 2009, pp. 214-215.

[3] Hoskote, Y., S. Vangal, A. Singh, N. Borkar, and S. Borkar (2007) "A 5-GHz Mesh Interconnect for a Teraflops Processor," Micro, IEEE, 27(5), pp. 51–61.

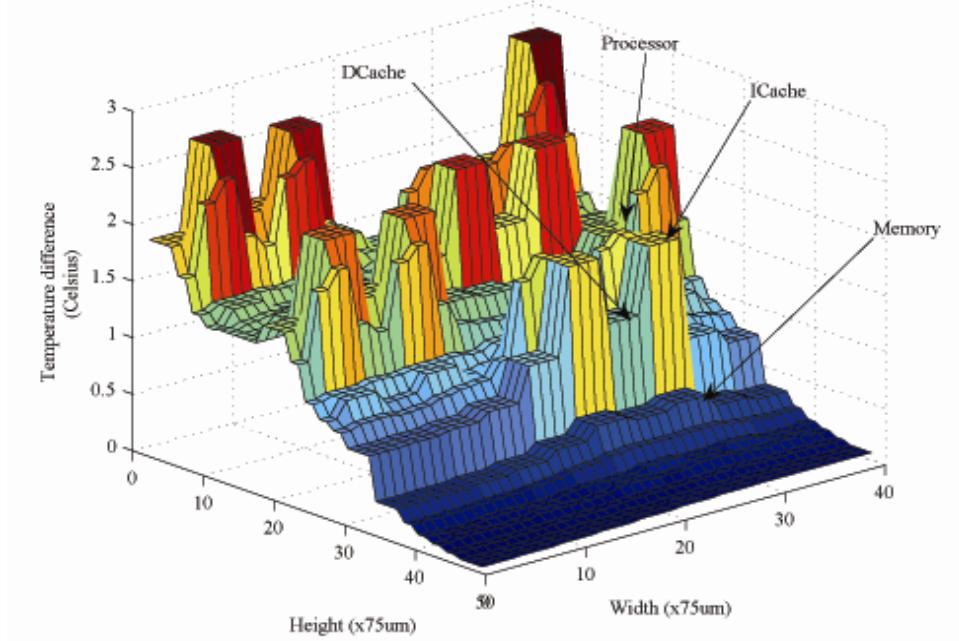
Hot spots and Thermal problems



Chip floorplan

Some hot spots in steady state:

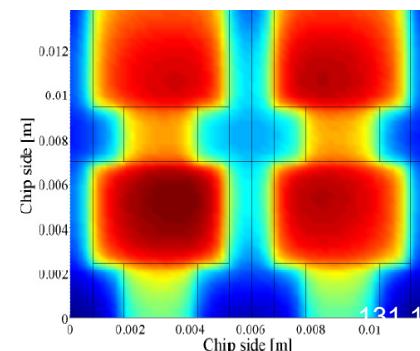
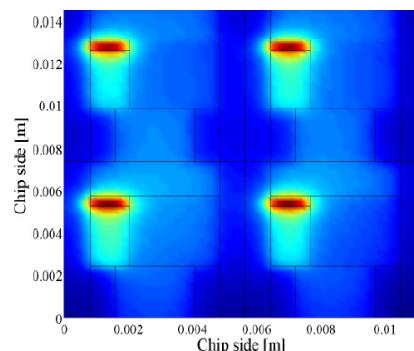
- Silicon is a good thermal conductor (only 4x worse than Cu) and temperature gradients are likely to occur on large dies
- Lower power density than on a high performance CPU (lower frequency and less complex HW)



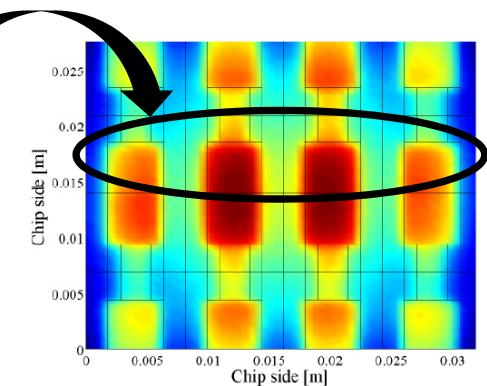
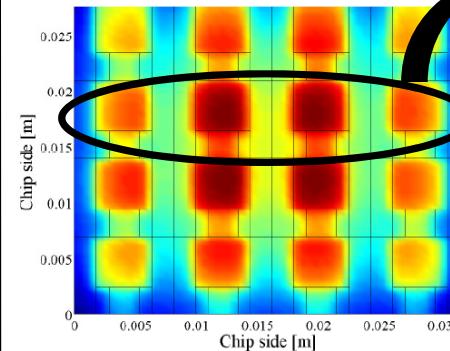
Steady state temperature

Thermal oriented analysis and design

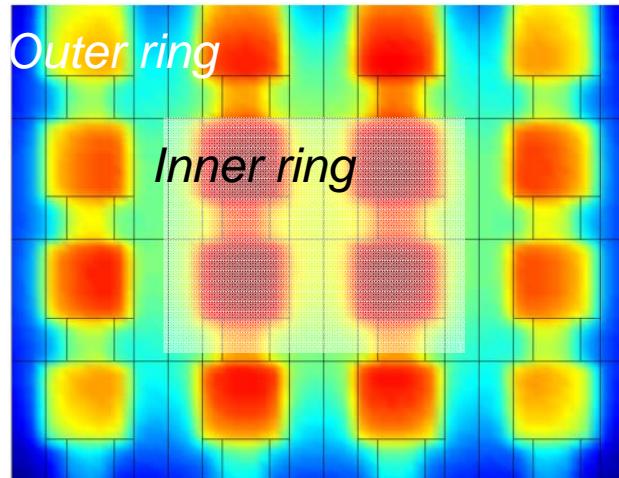
Multiple levels of detail



Different floorplan options



Practical for assess methodologies, for example



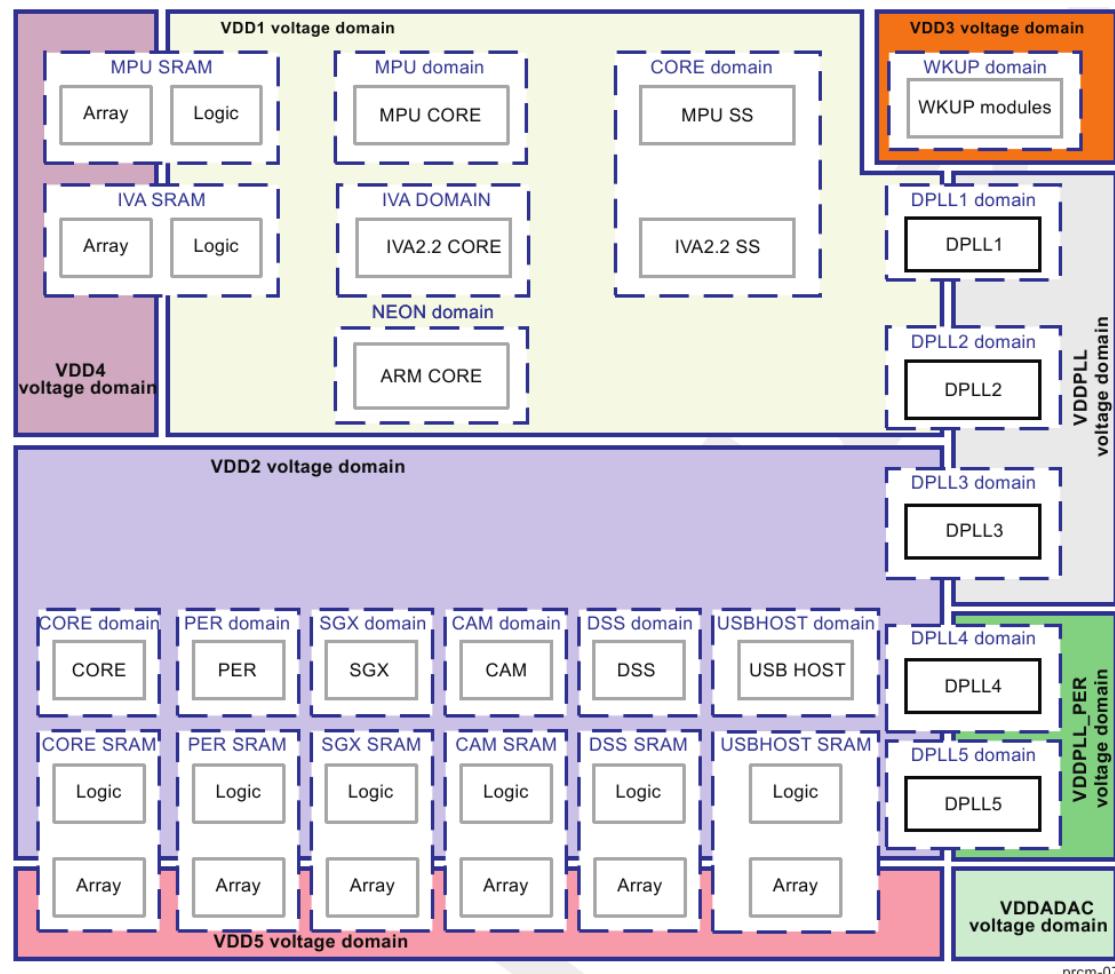
Goal: Fairly balance the chip thermal map

How to:

- Divide the chip in concentric rings, where a DVFS module can set frequency (f) and voltage (V)
- Collect experimental data to design-time optimize (V, f) pair values for each ring subject to
- Minimum difference f difference between each rings pair.
(HINT frequency proportional to performance and we want fairness)

Voltage and frequency scaling

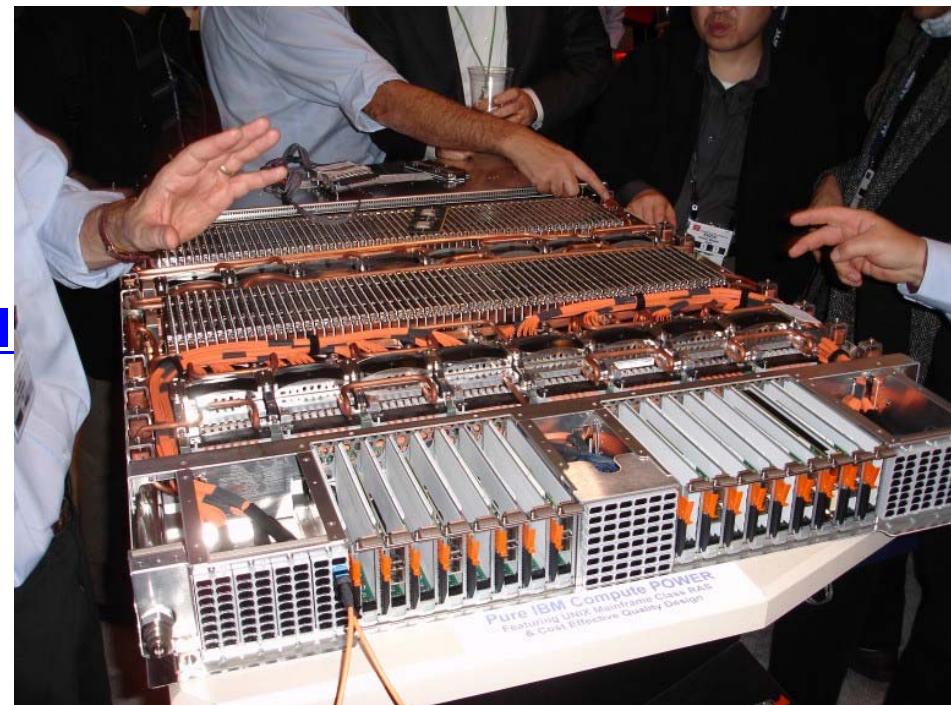
- Chips/Systems are partitioned in islands differing in terms of voltage and frequency, with the possibility to be switched off dynamically (power gating)



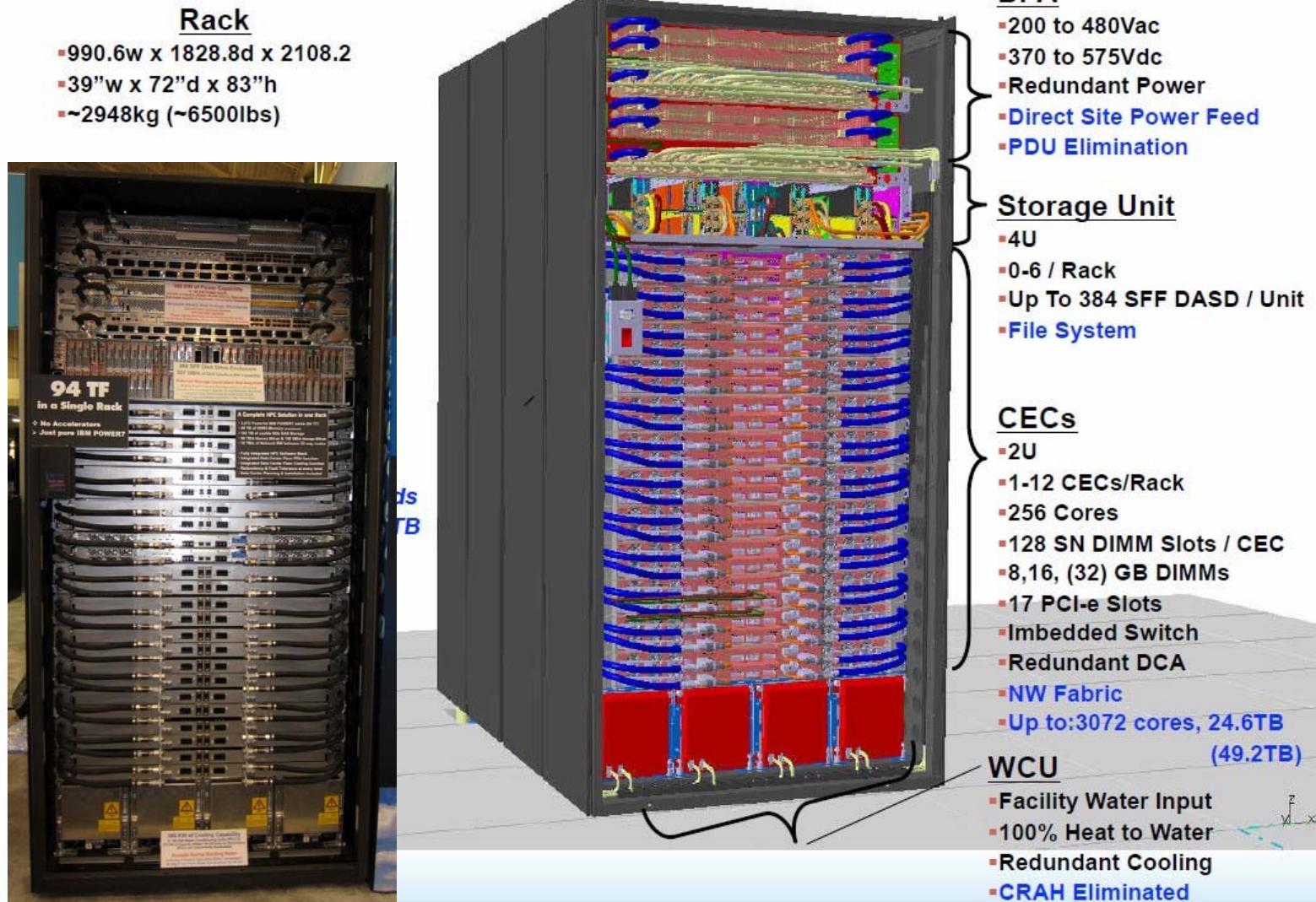
- OMAP Platform by Texas Instruments

Example: blue waters by IBM

- 10 PFlop (10^{16}) peak performance
- 300'000 compute cores = 37'500 CPU chips = 9375 QCM = 1172 drawers = 98 racks
- 800W / QCM 7.5 MW in CPUs
- New building completed
- 24 transformers@2 MW
- <http://www.ncsa.illinois.edu/Bl>



Blue Waters (cont'd)



Overview of HPC Cooling Systems

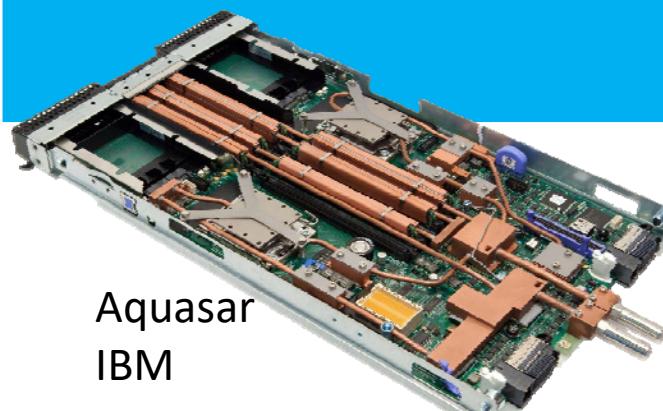
Air cooling

- Very low HTC
- Very low chip uniformity
- Large heat sinks
- Multiple air ducts in Datacenter
- Noisy
- Expensive maintenance
- Complex air management



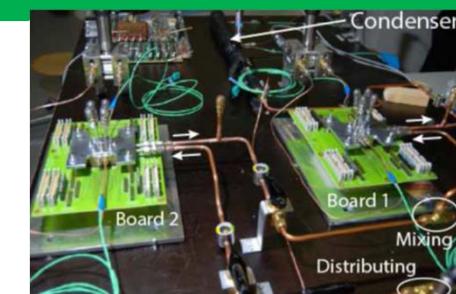
Water cooling

- + Less fans/ducts
- + Better HTC
- + Smaller heat sinks
- + Possible heat recovery
- Large Pumps

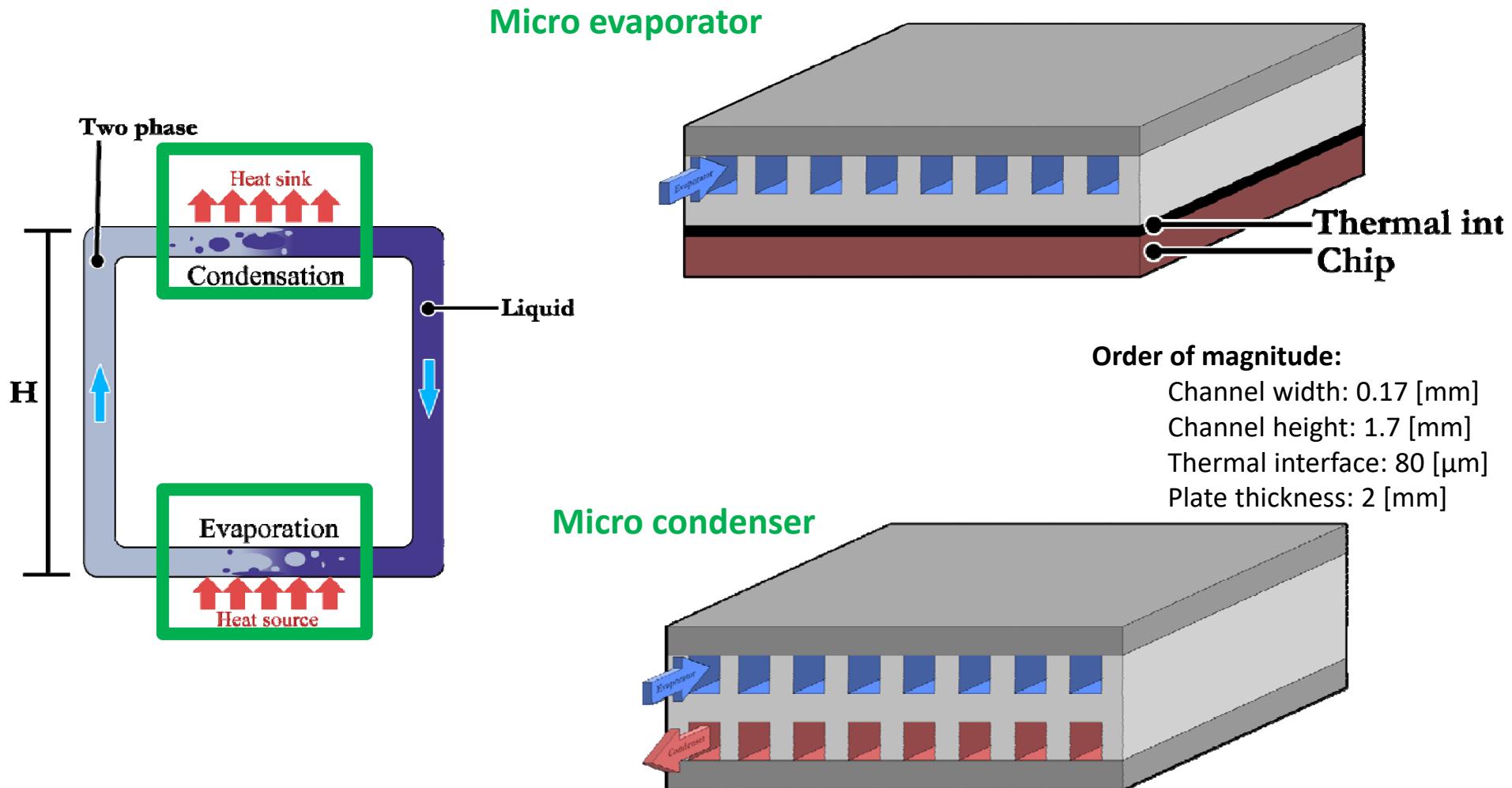


Two phase cooling

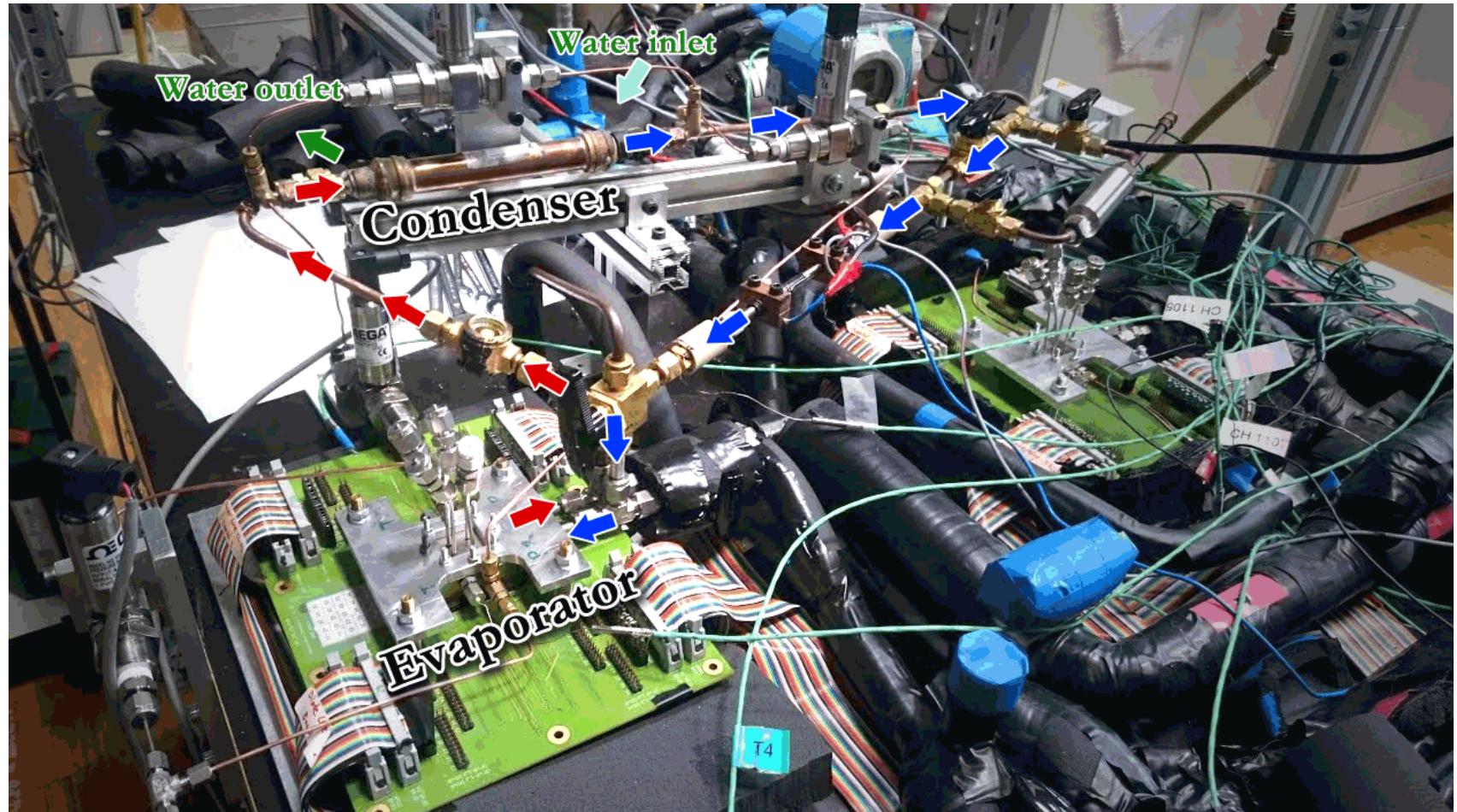
- + Smaller pump
- + Higher HTC
- + Better chip uniformity
- + Isothermal coolant
- + Good hot spot cooling
- + Possible heat recovery
- Low pump efficiency and reliability



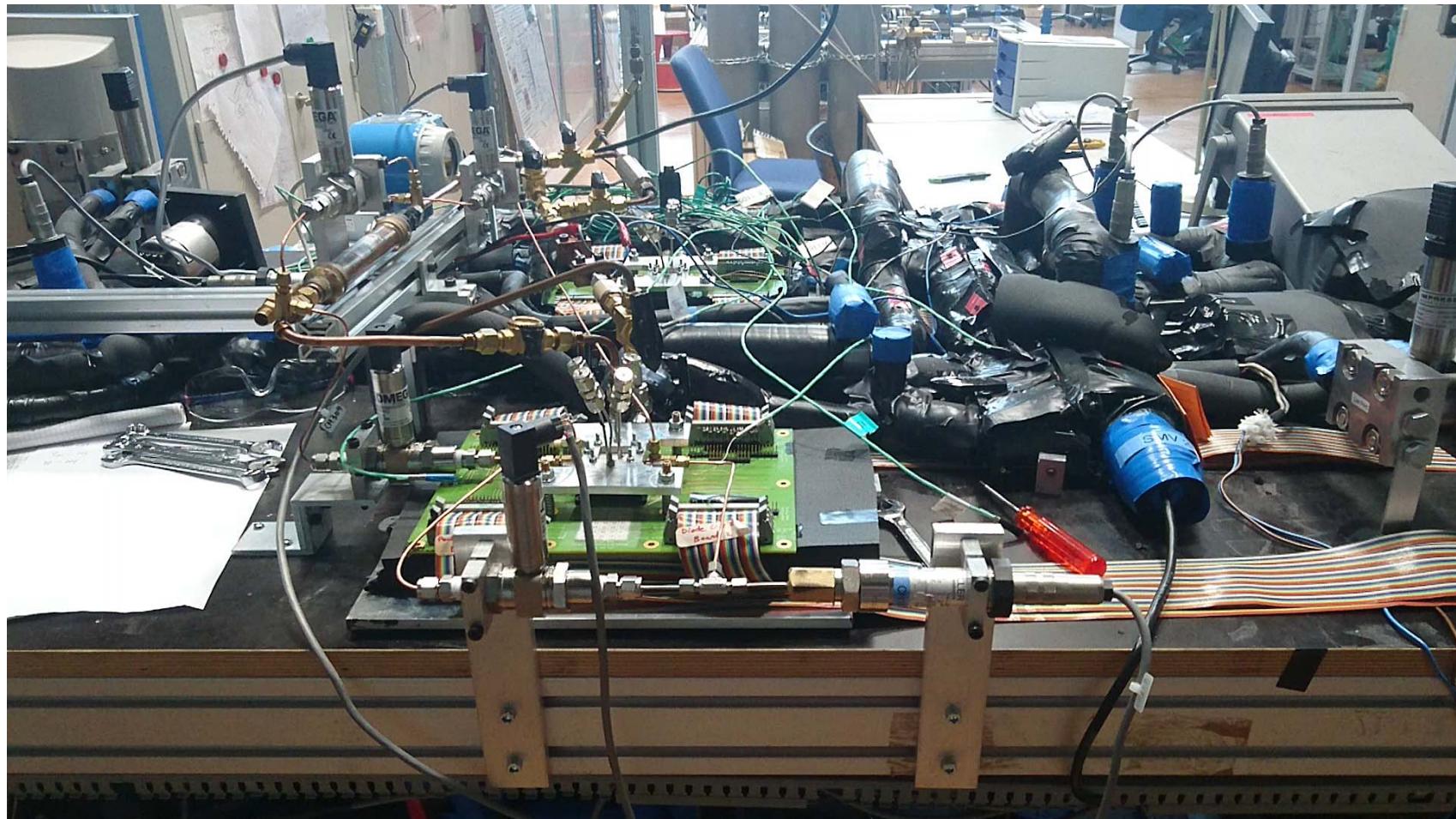
Proposed Cooling: Thermosyphon



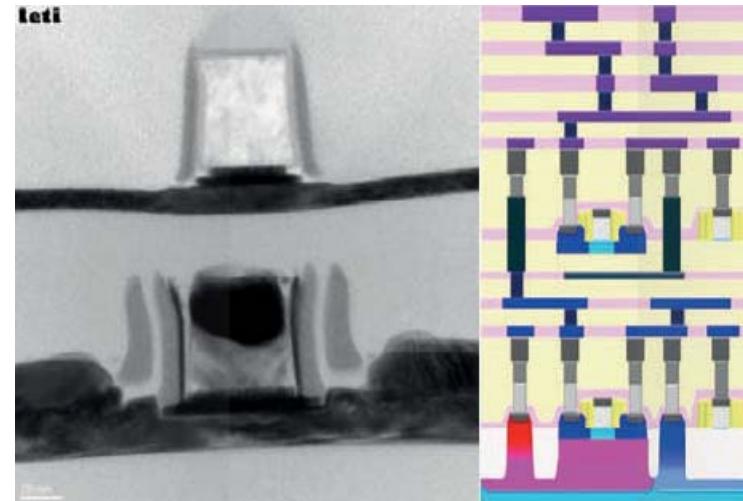
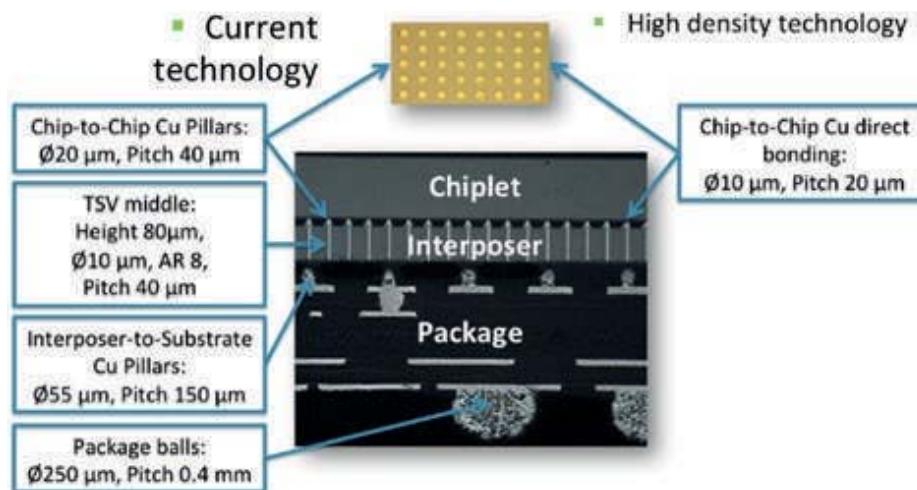
Thermosyphon Experim. Setup at EPFL



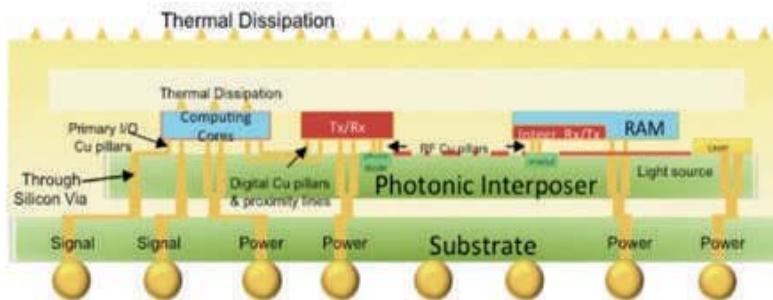
Thermosyphon Experim. Setup at EPFL



Entering into the third dimension

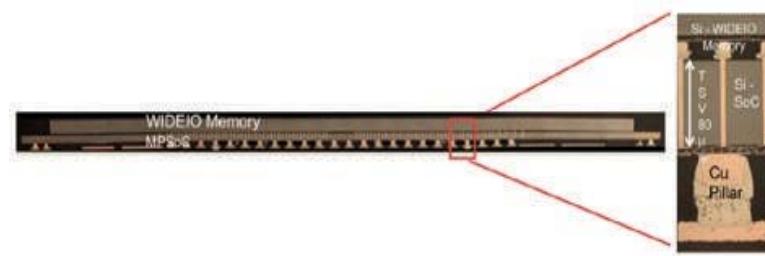


3D interconnect technology



Compute module using photonic interconnect between chiplets

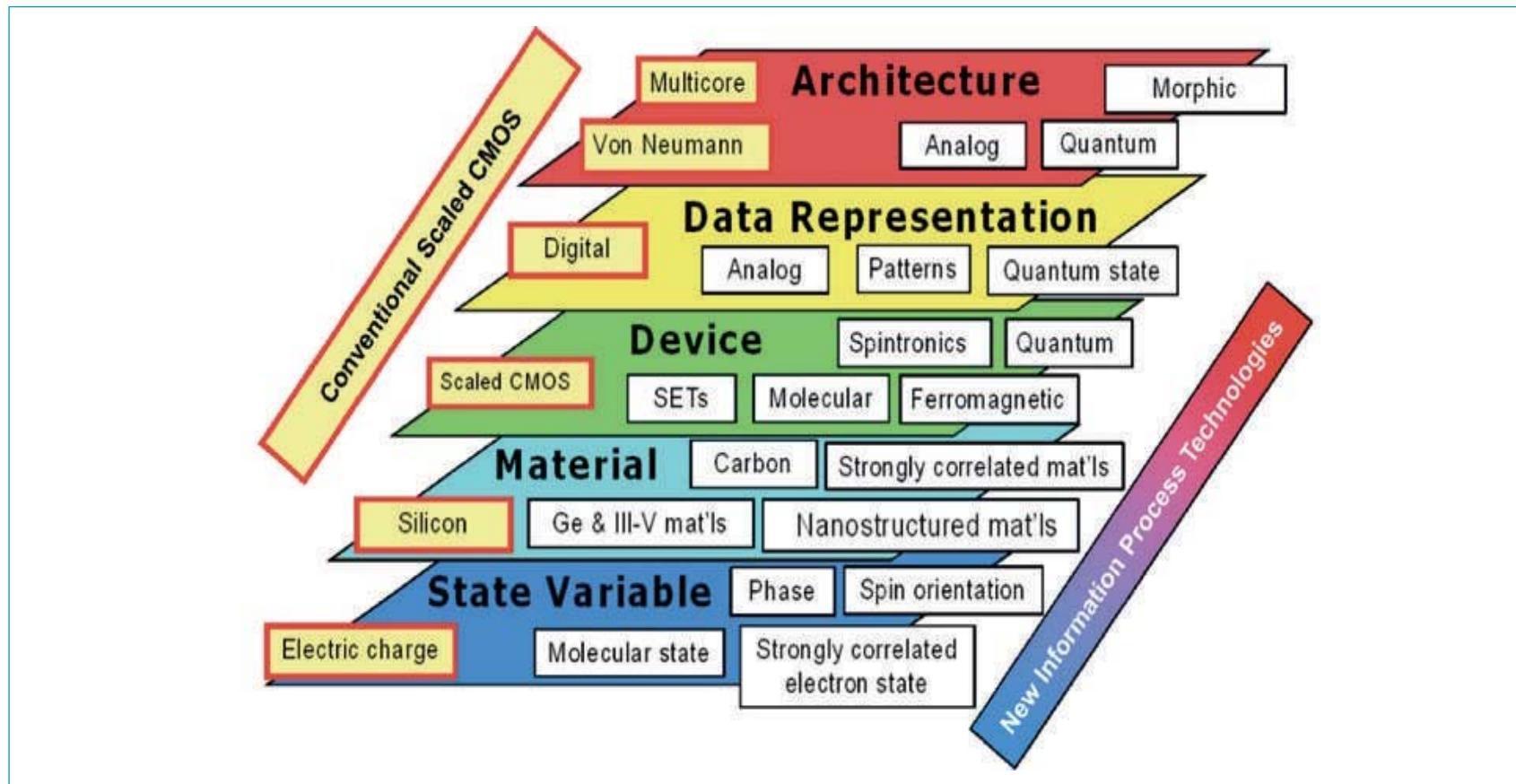
Monolithic 3D devices: transistors are built on top of each other



A Wide I/O memory stacked on top of a MPSoC

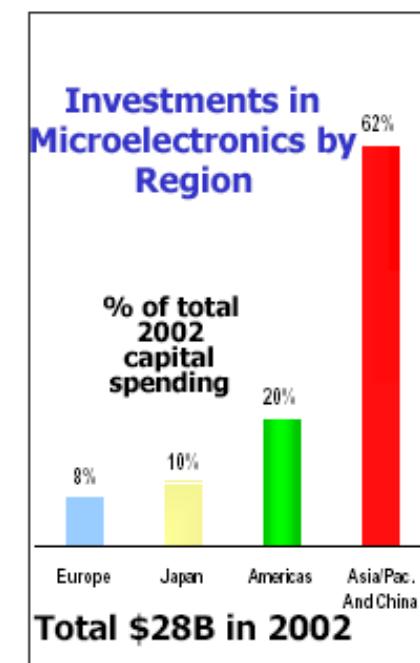
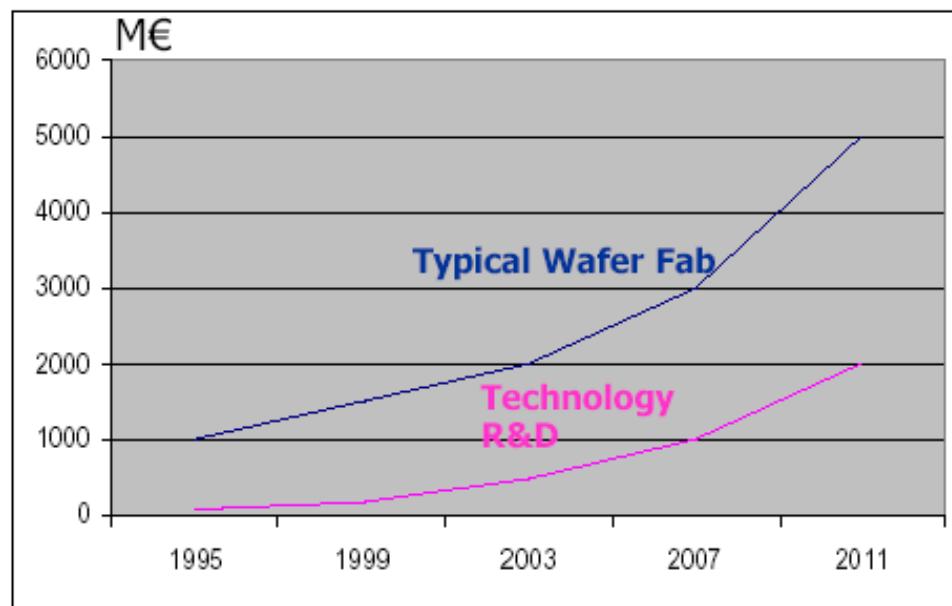
Do not forget...memory device technology

- Potential performance stagnation in the coming years?



Leadership and Competitiveness

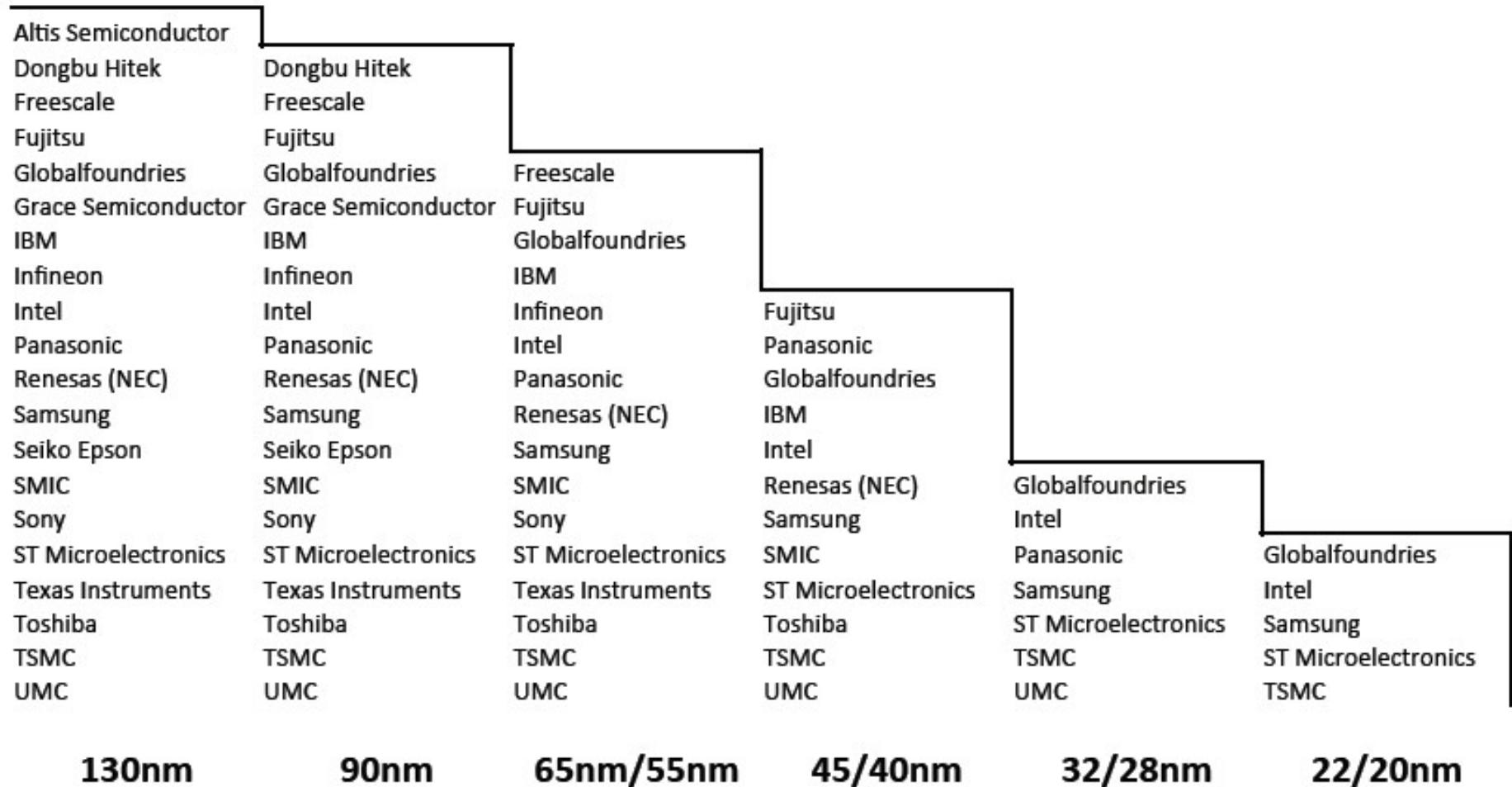
- Europe needs to quickly fill the gap on IP architectures and Computer Science
- Maybe is too late...



The Twilight of Moore's Law: Economics

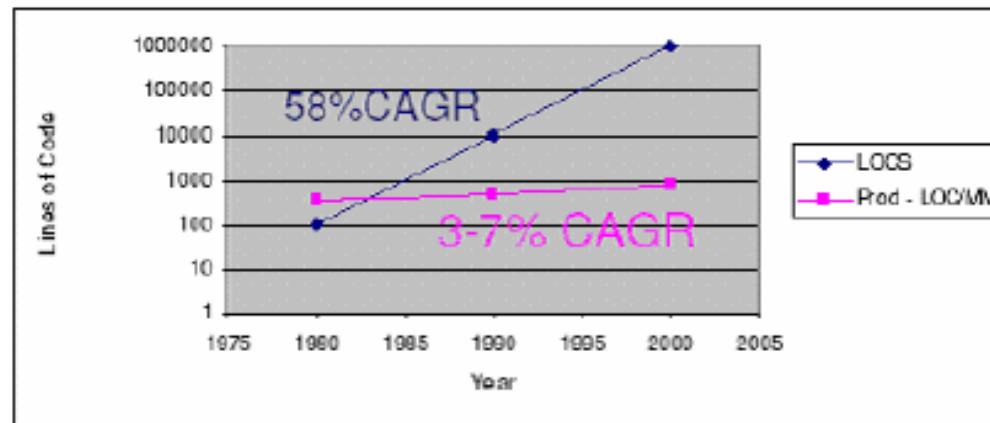
▪ Market volume wall

- only the largest volume products will be manufactured with the most advanced technology



Complexity vs productivity growth

SW complexity & productivity growth



HW complexity & productivity growth

