



# HiPEAC Network of Excellence

## HiPEAC vision 2015

Archi'15

Marc Durantou  
June 12<sup>th</sup>, 2015



# What is HiPEAC?

- HiPEAC is a European Network of Excellence on **H**igh **P**erformance and **E**Embedded **A**rchitecture and **C**ompilation
- Created in 2004, **HiPEAC** gathers over 370 leading European academic and industrial computing system researchers from nearly 140 universities and 70 companies in one virtual centre of excellence of 1500 researchers.



- Coordinator: Koen De Bosschere (UGent)



# HiPEAC mission:

HiPEAC encourages computing innovation in Europe by providing:

- collaboration grants, internships, sabbaticals,
- the semi-annual computing systems week,
- The ACACES summer school,
- the yearly HiPEAC conference.



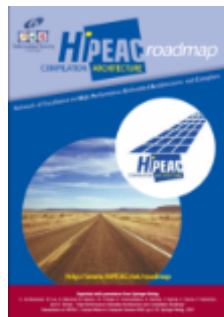
The 10th HiPEAC conference took place in Amsterdam, The Netherlands, January 19-21, 2015 and gathers more than 600 people

# The HiPEAC Vision



New HiPEAC Vision Document was published in January 2015

**Editors: Marc Duranton** (FR-CEA), **Koen de Bosschere** (BE-U Gent), **Albert Cohen** (FR-INRIA), **Jonas Maebe** (BE-U Gent), **Harm Munk** (NL-ASTRON)



2008



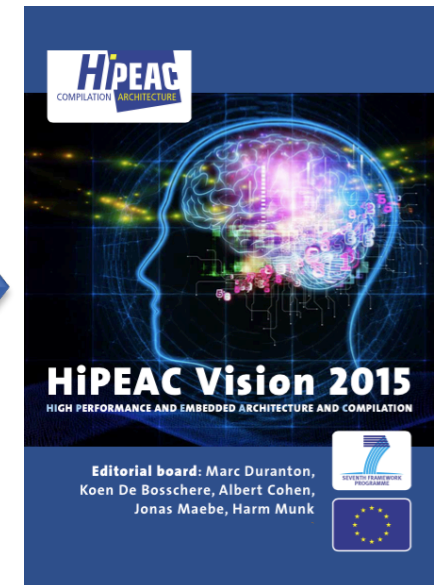
2009



2011



2013



2015

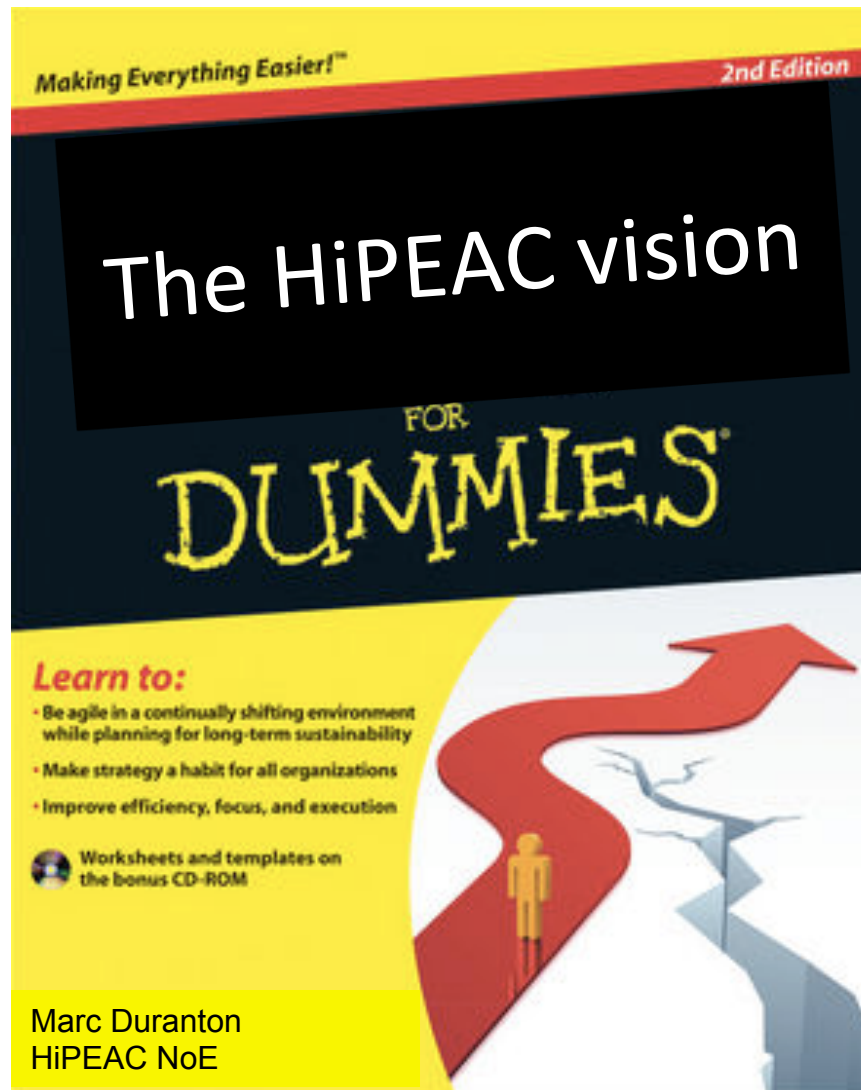
<http://www.hipeac.org/vision/>



# Warning !!!



## *Few technical words inside !!!*





## *On the road for the HiPEAC Vision 2015*

The document is based on the inputs from:

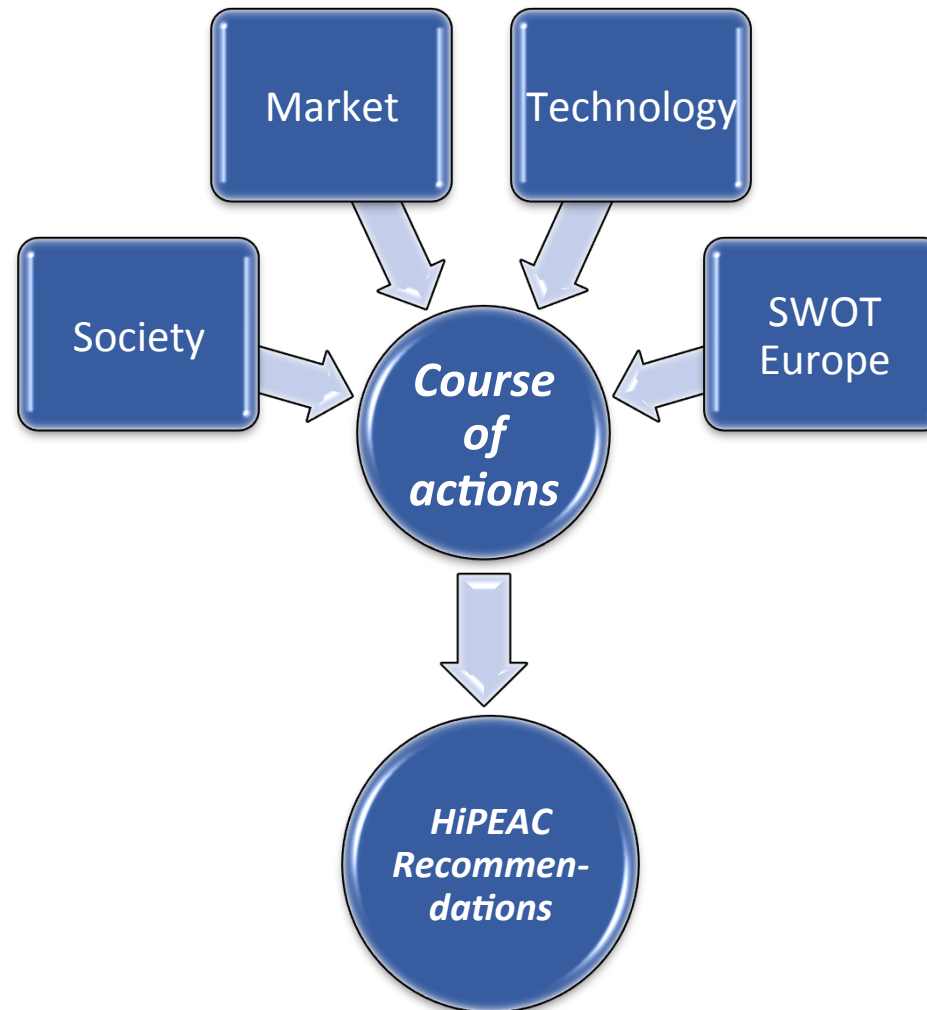
- The HiPEAC members
  - Survey sent to all members, more than 50 feedbacks
  - Dedicated session the the CSW of May 2014 in Barcelona
  - Feedback session during the Autumn CSW in Athens
  - Free participation and feedback from all partners and members throughout the whole process
- The teachers of the ACACES summer schools 2013 and 2014
  - Evening session with teachers and industrials



## *On the road for the HiPEAC Vision 2015*

- Two HiPEAC organized workshops with HiPEAC members and external invitees
  - *“New computing approaches and approximate computing”*, 27 May 2014, Brussels.
  - A meeting on more general topics, 13 June 2014, Brussels.
- Three workshops organised in cooperation with the DGConnect - Complex Systems and Advanced Computing:
  - *“Next Generation Computing Systems: components and architectures for a scalable market”*, 10 Dec, 2013, Brussels.
  - *“Software tools for next generation computing”*, 24 June 2014, Brussels.
  - *“Energy-Efficient Computing Systems, dynamic adaptation of Quality of Service and approximate computing”*, 27 Nov. 2014, Brussels,

# Structure of the HiPEAC vision 2015







## Highlights of the **HiPEAC Vision 2015**

For the first time, we have noticed that the community really *starts looking for disruptive solutions,*  
*and that incrementally improving current technologies is considered inadequate to address the challenges that the computing community faces:*

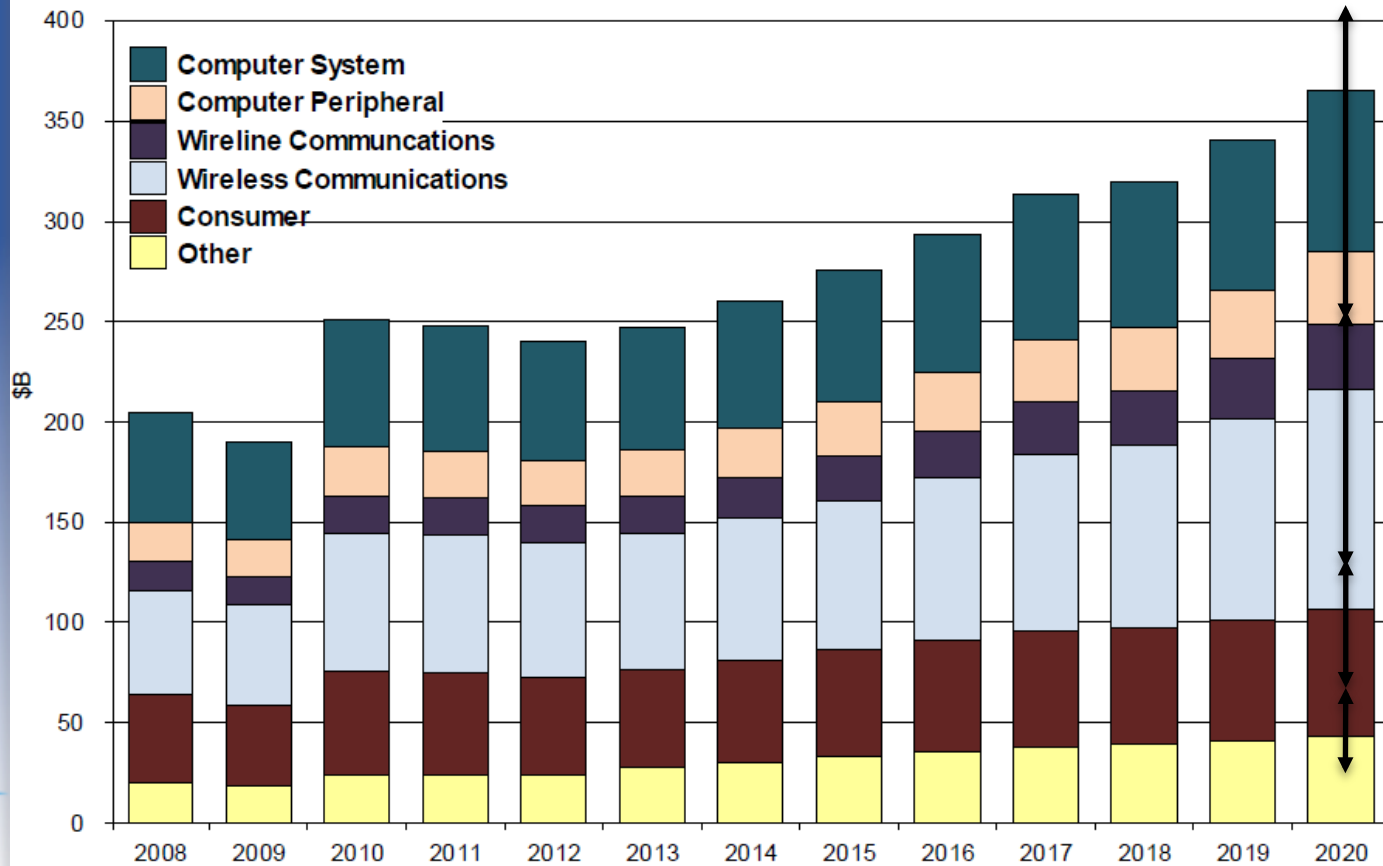
***“The End of the World  
As We Know It”***



# IC Market (source IBS)

- IC market by 2020: \$ 350 B
- 1/3 of total IC market for computing applications
- 1/3 mobile/IoT, 1/3 automotive/industrial

IC Market by Application Segment



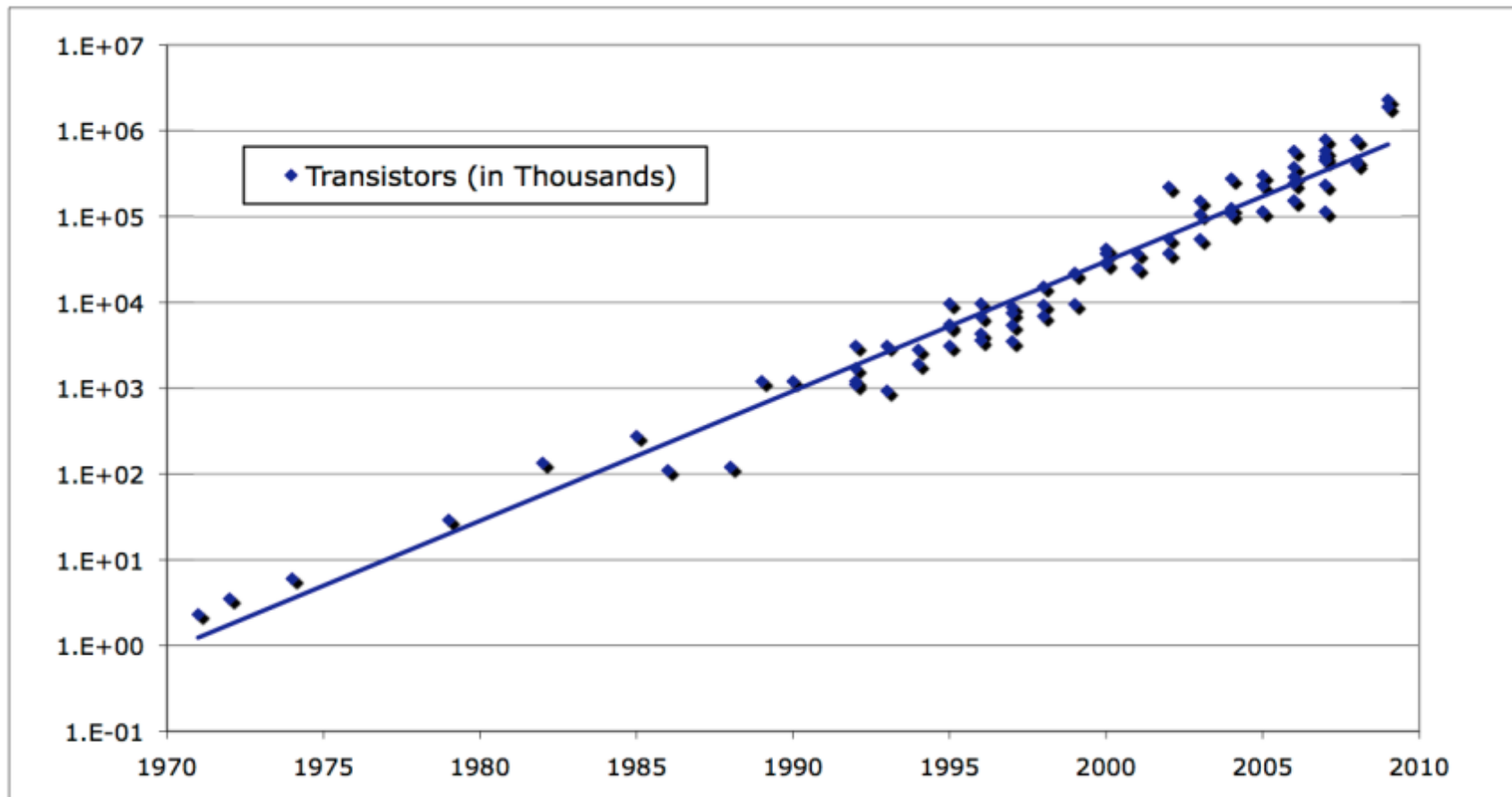
Processors  
(Server, PC...)  
Networking,  
DRAM, Flash.

Mobile phone,  
base station,  
IoT.

Console, camera,  
audio...

Automotive,  
industrial and  
medical

# Moore's law: increase in transistor density



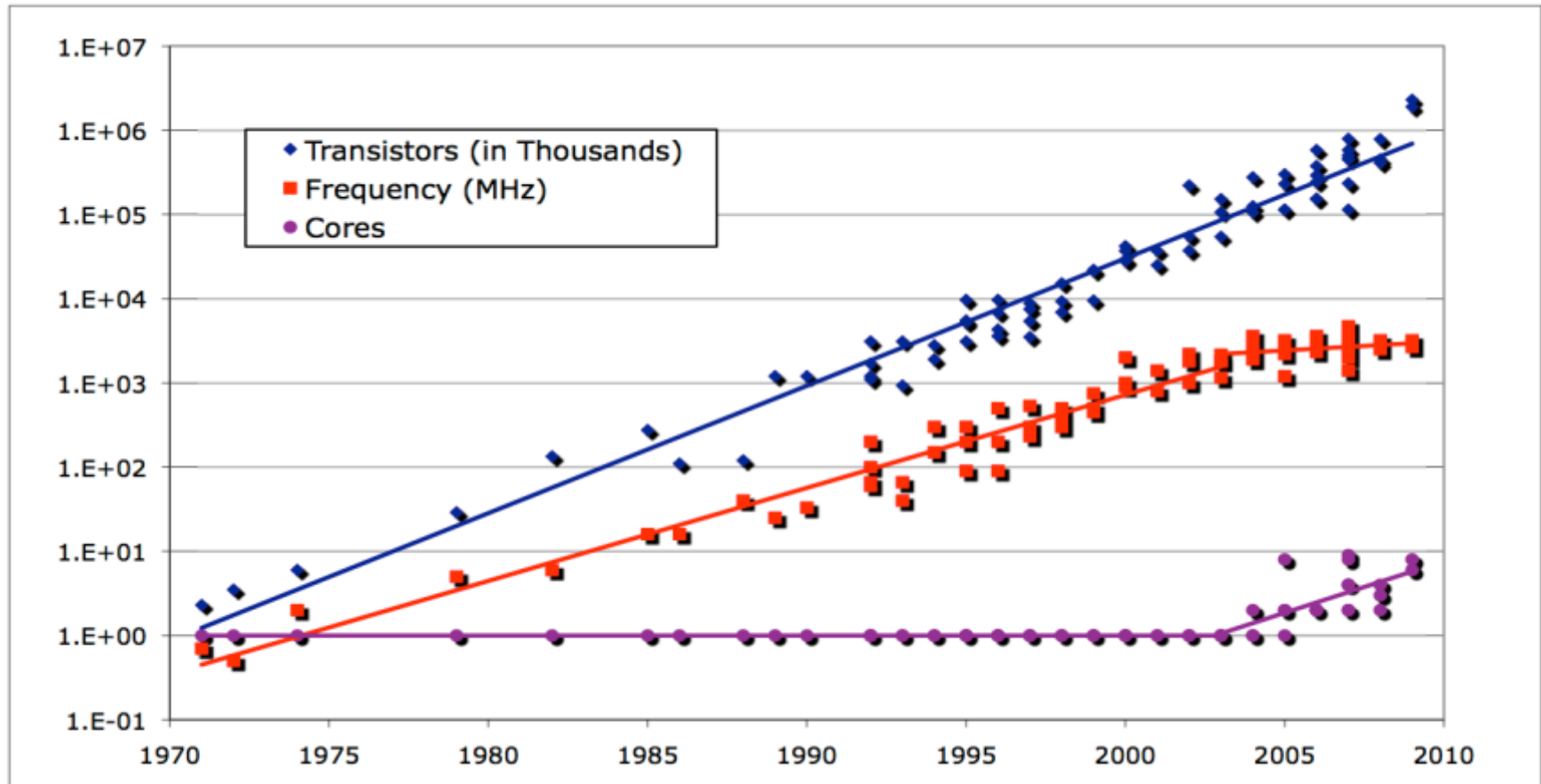
Source from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović

# The end of Dennard Scaling

Parameter (scale factor = a)	Classic Scaling	<p>Everything was easy:</p> <ul style="list-style-type: none"><li>• Wait for the next technology node</li><li>• Increase frequency</li><li>• Decrease V<sub>dd</sub></li></ul> <p>⇒ Similar increase of sequential performance ⇒ No need to recompile (except if architectural improvements)</p>
Dimensions	1/a	
Voltage	1/a	
Current	1/a	
Capacitance	1/a	
Power/Circuit	1/a <sup>2</sup>	
Power Density	↓	
Delay/Circuit	1/a	

Source: Krisztián Flautner “From niche to mainstream: can critical systems make the transition?”

# Limited frequency increase $\Rightarrow$ more cores



Source from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović

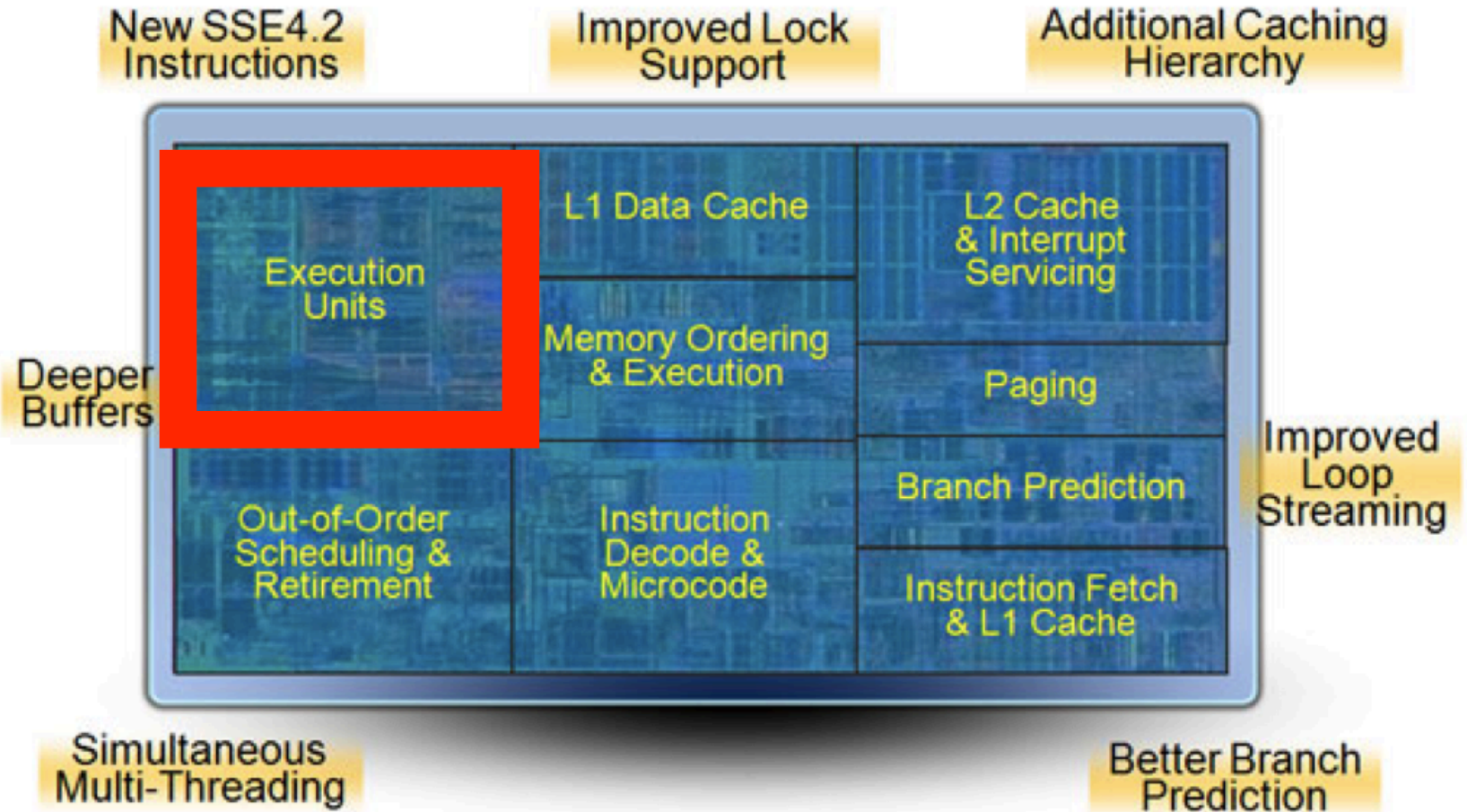


# Why using several compute cores?

1. Using several cores is also an answer to the Law of Diminishing Returns [Pollack's Rule] :
  - Effectiveness per transistor decreases when the size of a single core is increased, due to the locality of computation
  - Controlling a larger core and data transport over a single larger core is super-linear
  - **Smaller cores are more efficient** in ops/mm<sup>2</sup>/W
2. Large area of today's microprocessors are for best effort processing and used to cope with unpredictability (branch prediction, reordering buffers, instructions, caches).



# Less than 20% of the area for execution units

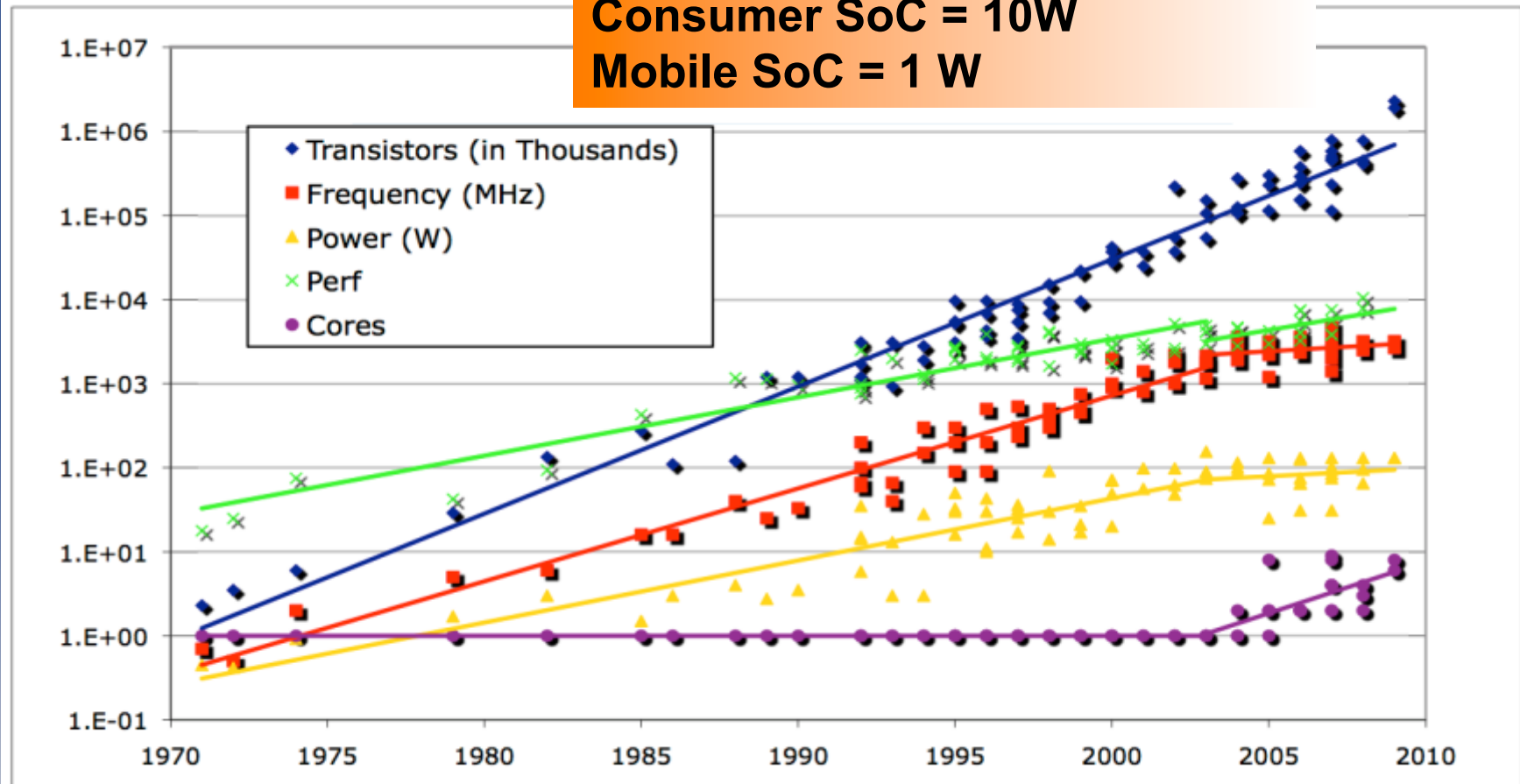


Source: Dan Connors, "OpenCL and CUDA Programming for Multicore and GPU Architectures» ACACES 2011



# Limitation by power density and dissipation

**GP CPU = 200 W (45 nm)**  
**Consumer SoC = 10W**  
**Mobile SoC = 1 W**



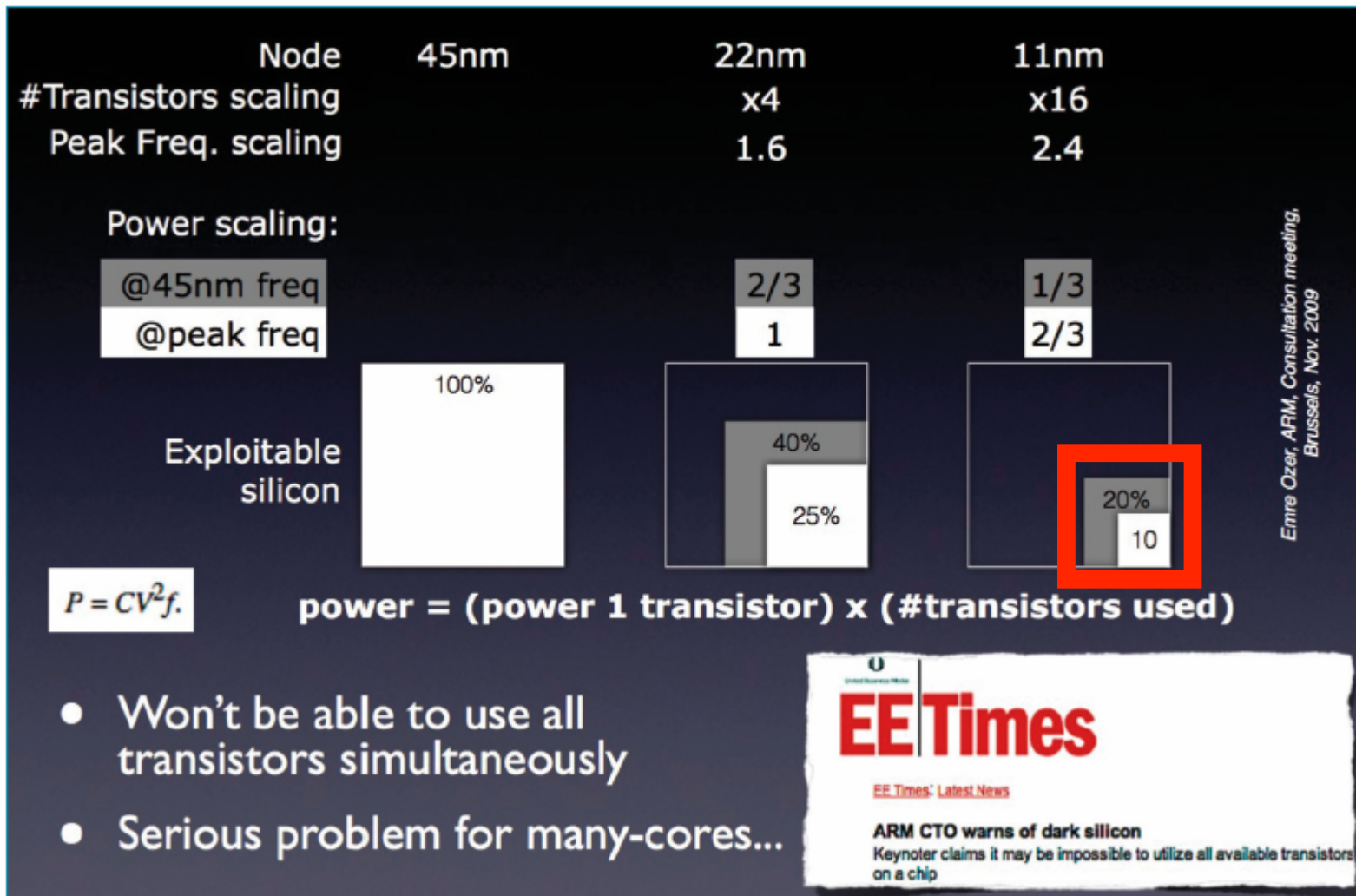
Source from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanovic



# Power limits the active silicon area

=> “Dark silicon”

=> More efficient specialized units



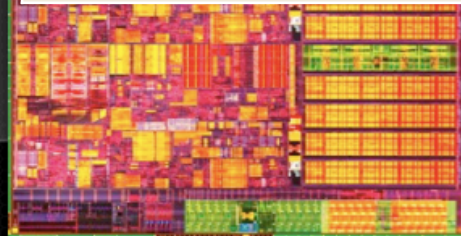
Emre Ozer, ARM, Consultation meeting, Brussels, Nov. 2009

# Specialization leads to more efficiency

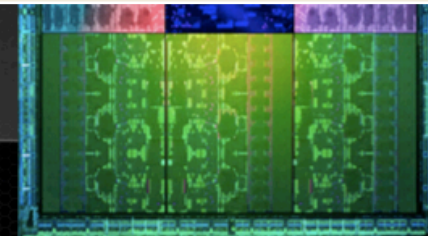
**CPU**  
1690 pJ/flop

**GPU**  
140 pJ/flop

Type of device	Energy / Operation
CPU	1690 pJ
GPU	140 pJ
Fixed function	10 pJ



Westmere  
32 nm



Kepler  
28 nm

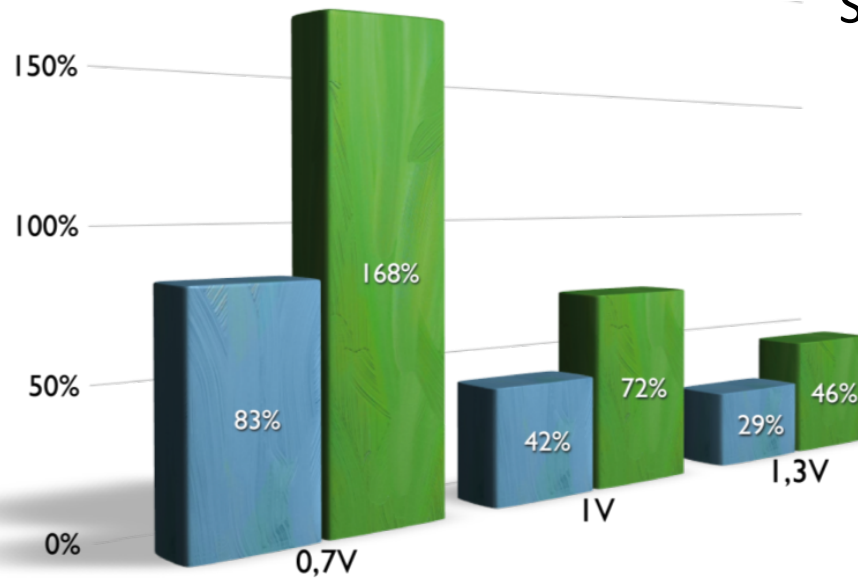
Source from Bill Dally (nVidia) « Challenges for Future Computing Systems »  
HiPEAC conference 2015



# Energy efficient technology: FDSOI

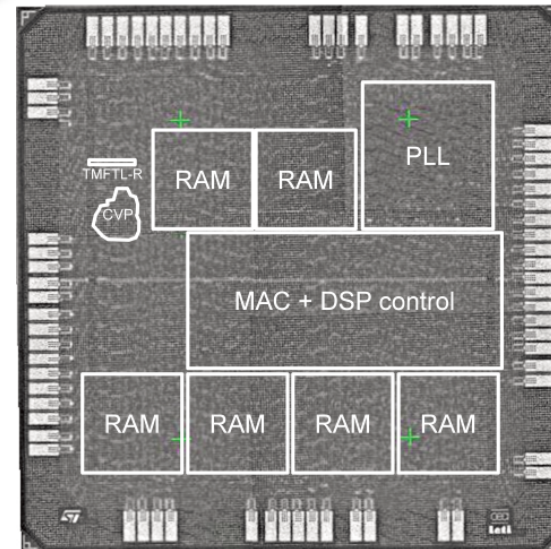
UTBB-FDSOI performance gain versus conventional Bulk CMOS technology.

Blue: no body biasing, Green: FBB = +1V.



- Demonstrated by CEA tech and STMicroelectronics (ISSCC 2014)
- Ultra-Wide Voltage Range (UWVR) operations:  $VDD=[0.39V - 1.3V]$
- High-frequency:
  - Fclk > 2.6GHz @ 1.3V**
  - Fclk > 450MHz @ 0.39V**

- **Fully Depleted – Silicon on Insulator**
  - **Improved performance-per-watt**
  - **Adaption to variability of loads under software control**



Technology	UTBB FDSOI 28 nm
Transistors	Flip-Well (LVT) L=24nm
Core area	1 mm <sup>2</sup>
DSP benchmark	FFT 1024
VDD range	0.397V-1.3V
VBB range	0V±2V

Prototype chip from Beigné, ISSCC '14



# Potential other optimizations

$$P_{\text{per unit}} = C V^2 f + T_{\text{sc}} V I_{\text{peak}} + V I_{\text{leak}}$$

Average power, peak power, power density, energy-delay, ...

## CIRCUITS

- **Voltage scaling/islands**
- **Clock gating/routing**  
Clock-tree distribution, half-swing clocks
- **Redesigned latches/flip-flops**  
pin-ordering, gate restructuring, topology restructuring, balanced delay paths, optimized bit transactions
- **Redesigned memory cells**  
Low-power SRAM cells, reduced bit-line swing, multi-Vt, bit line/word line isolation/segmentation
- **Other optimizations**  
Transistor resizing, GALS, low-power logic

## ARCHITECTURE

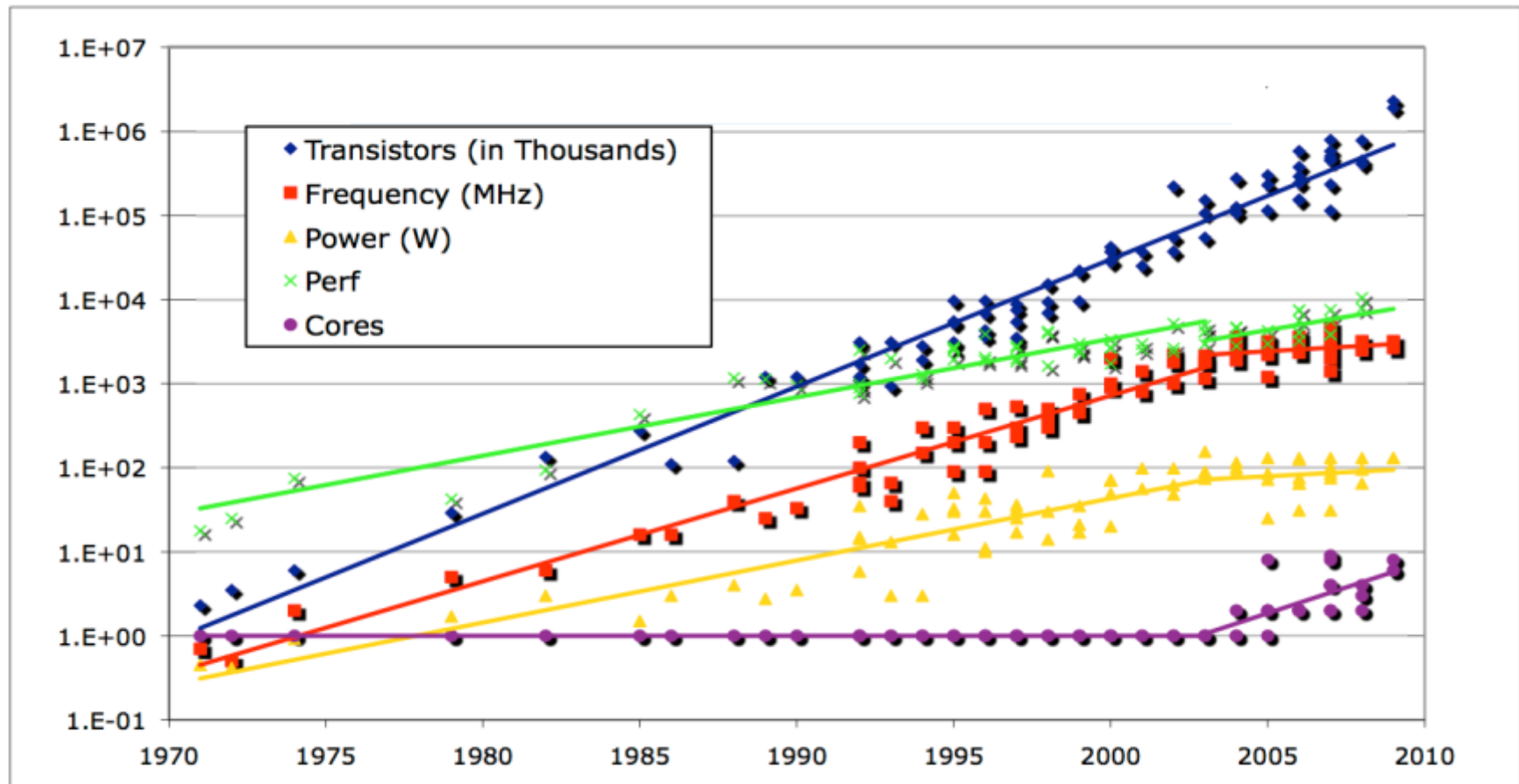
- **Voltage/freq scaling**
- **Gating**  
Pipeline, clock, functional units, branch prediction, data path
- **Split instructn windows**
- **SMT thread throttling**
- **Bank partitioning**
- **Cache redesign**  
Sequential, MRU, hash-rehash, column-associative, filter cache, sub-banking, divided word line, block buffers, multi-divided module, scratch
- **Low-power states**
- **DRAM refresh-control**
- **Switching control**  
Gray, bus-invert, address-increment
- **Code compression**
- **Data packing/buffering**

## COMPILER, OS, APPLICATION

- **Switching control**  
Register relabeling, operand swapping, instruction scheduling
- **Memory access reduce**  
Locality optimizations, register allocation
- **Power-mode-control**
- **CPU/resource schedule**
- **Memory/disk control**  
Disk spinning, page allocation, memory mapping, memory bank control
- **Networking**  
Power-aware routing, proximity-based routing, balancing hop count, ...
- **Distributed computing**  
Mobile agents placement, network-driven computation
- **Fidelity control**
- **Dynamic data types**
- **Power API**

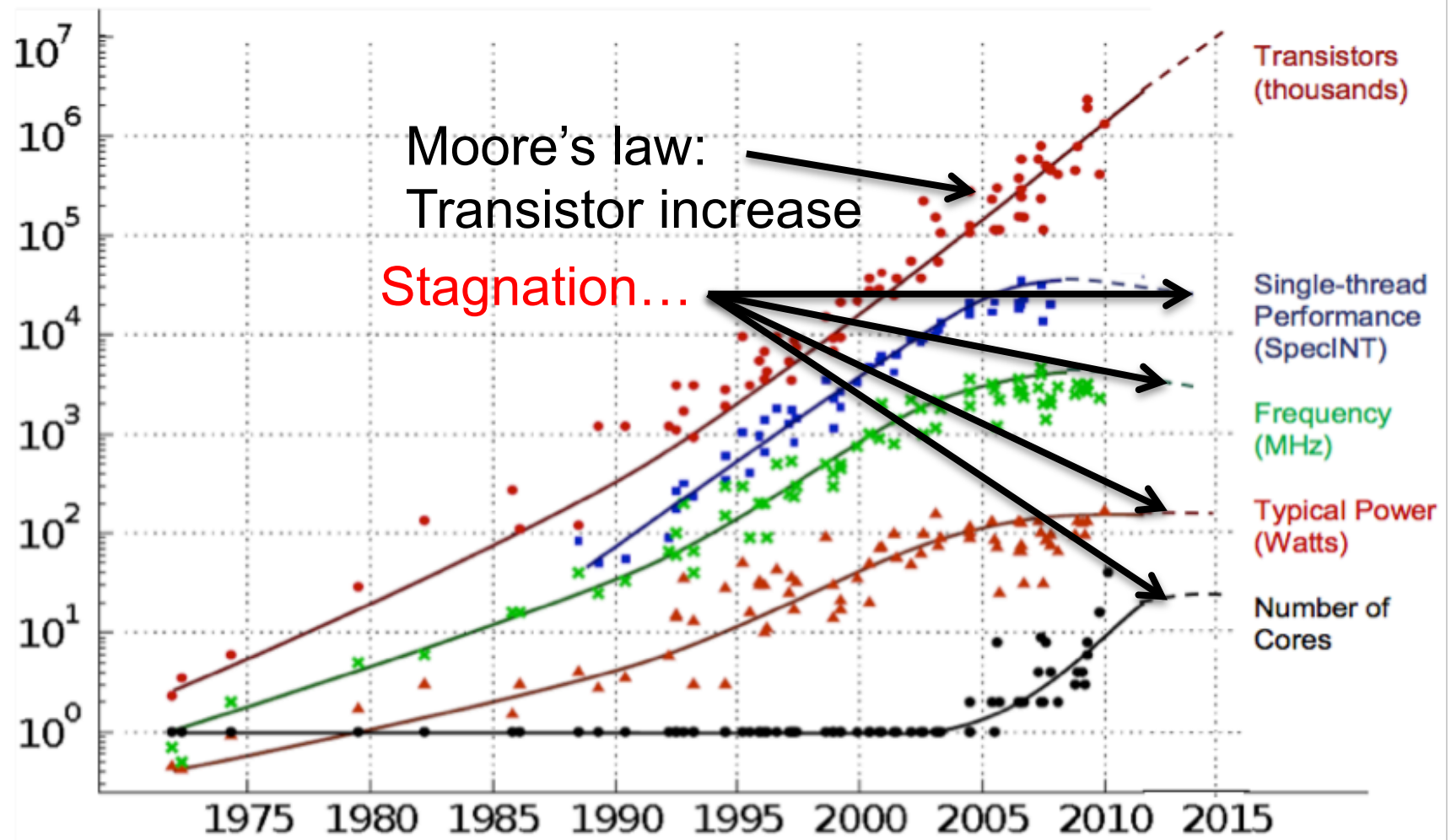
Source: P. Ranganathan, "System architectures for servers and datacenters »

# Limitation by power density and dissipation



Source from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanovic

# Stagnation of performance since few years

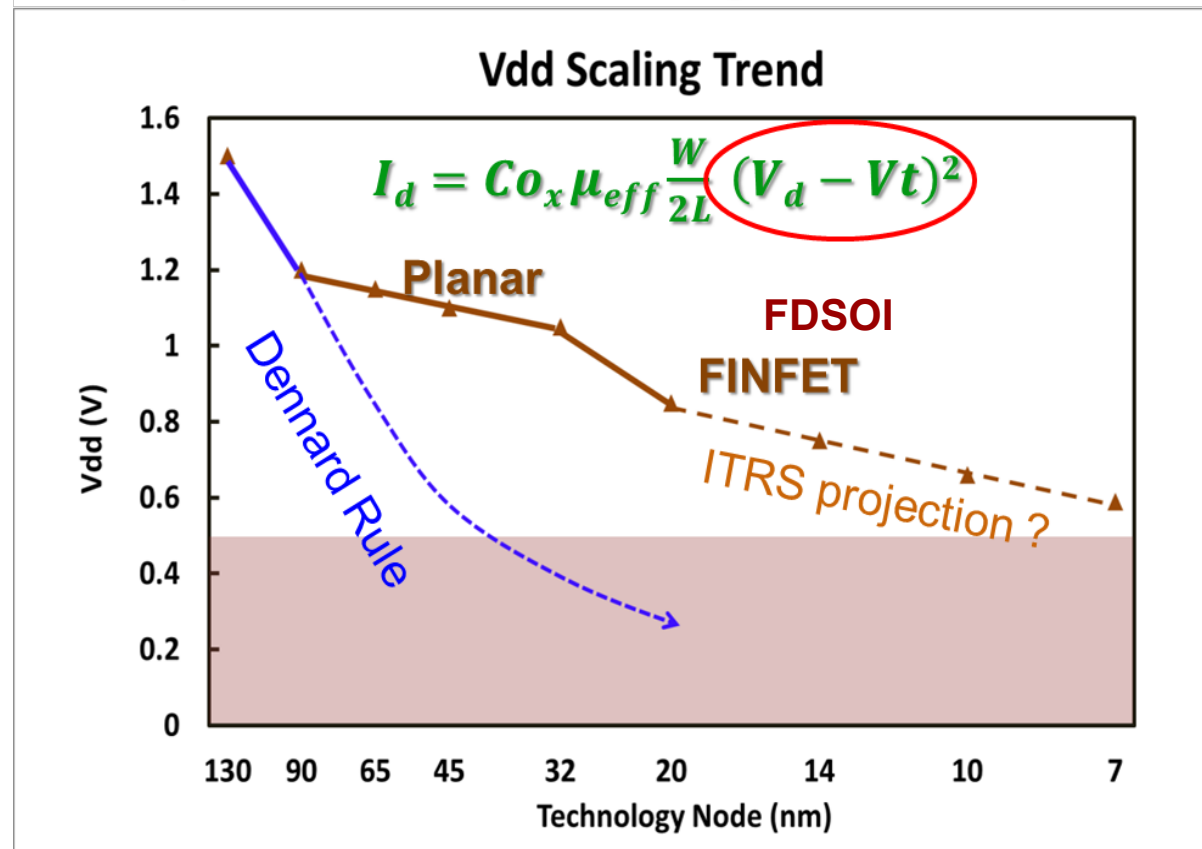


Source from C Moore, « Data Processing in ExaScale-Class Computer Systems », Salishan, April 2011

# Vdd scaling slowed down since N90

## Mobile Computing – The Vdd Scaling Issue

Vdd scaling most difficult:  
 $V_t$  associated with leakage – SCEs, RDF

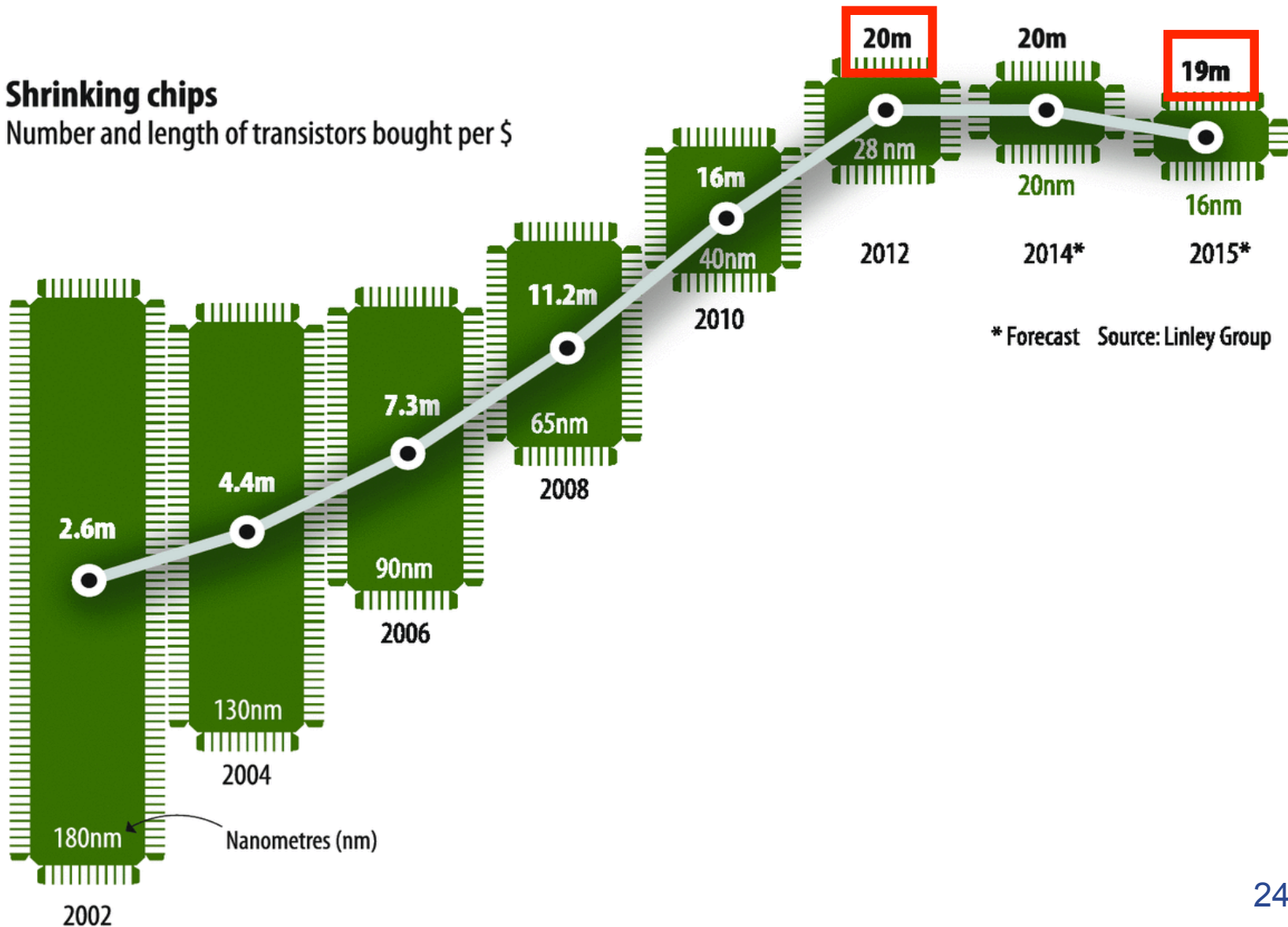


*Low Power Device: SS, DIBL,  $\sigma\Delta V_t$*

# The cost per transistors is not decreasing anymore

## Shrinking chips

Number and length of transistors bought per \$

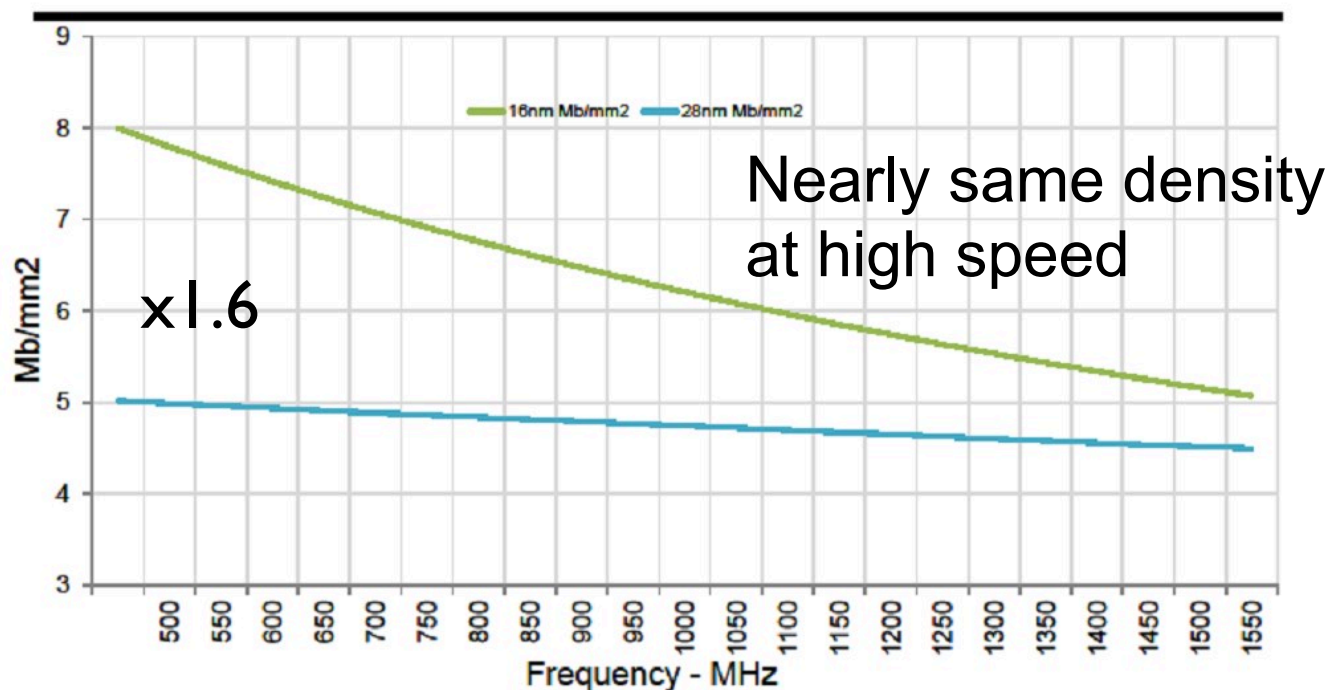




# Performances of SRAM hardly increase

Node	45nm	16nm	14nm	10 nm
Density	150 F <sup>2</sup>	237 F <sup>2</sup>	300 F <sup>2</sup>	450 F <sup>2</sup>

SRAM DENSITY - 16nm vs 28nm



Memory density at 1500MHz and above scales by ~1.1x or less from 28nm to 16nm

Source: Joel Hruska, « Stop obsessing over transistor counts: It's a terrible way of comparing chips », <http://www.extremetech.com>

# SRAM takes more and more SoC area

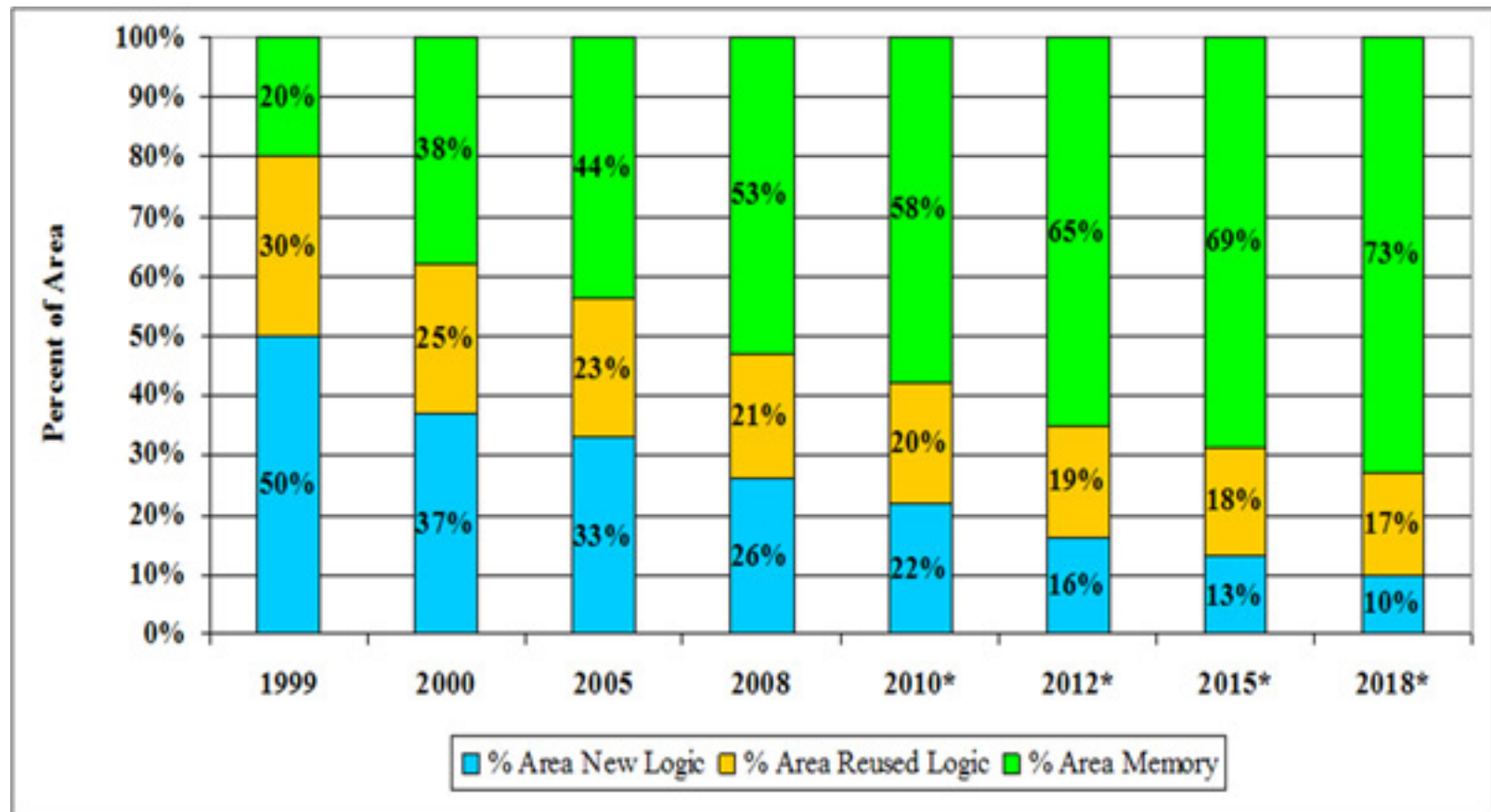


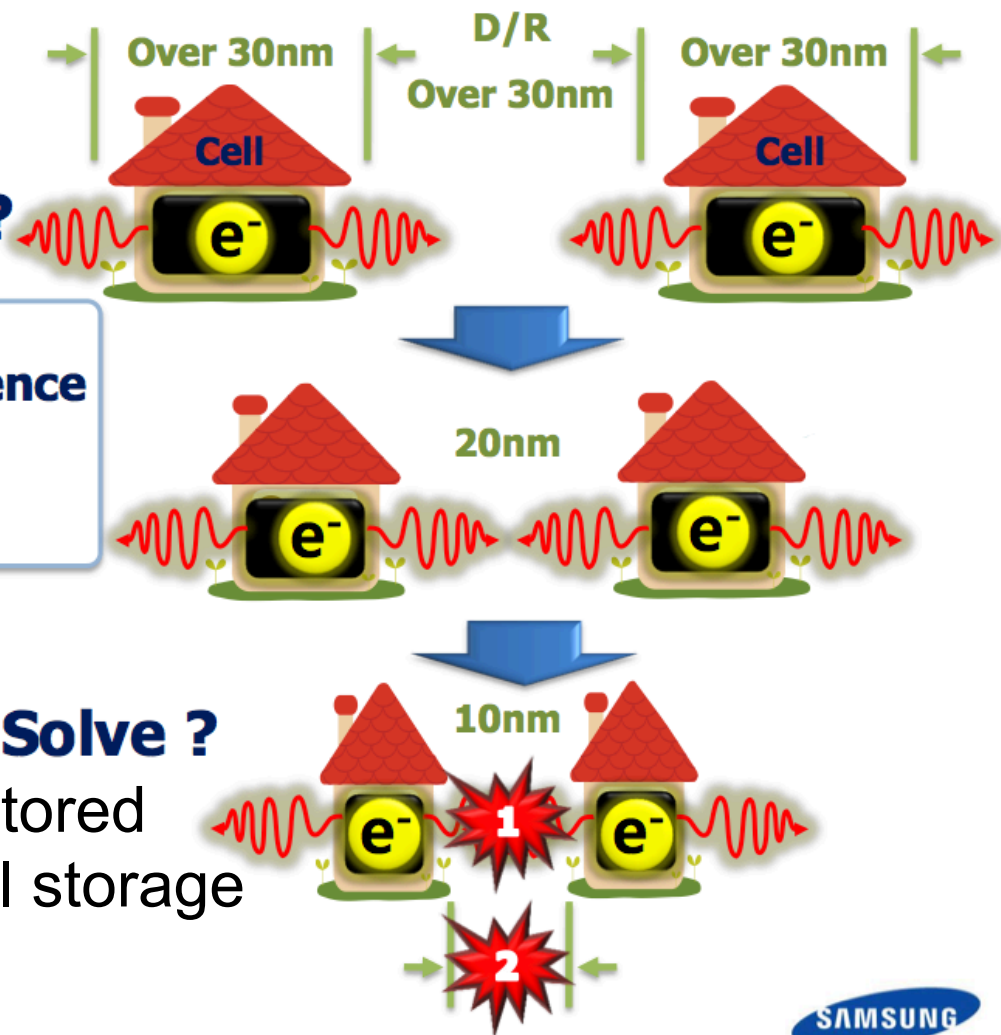
Fig 3. To compress design schedule time, designers often reuse earlier design blocks and use third party IP. It is very rare that a new chip featuring billions of transistors is designed completely from scratch. Generally, most of a new design's transistors are used to form memories or functions derived from similar functions implemented in earlier designs. (Source: Semico Research Corporation, Study Number SC103-10, October 2010)

# Flash scaling also hits limits

## 2 Questions

Q1. Why so difficult ?

- 1 Cell to Cell Interference
- 2 Patterning

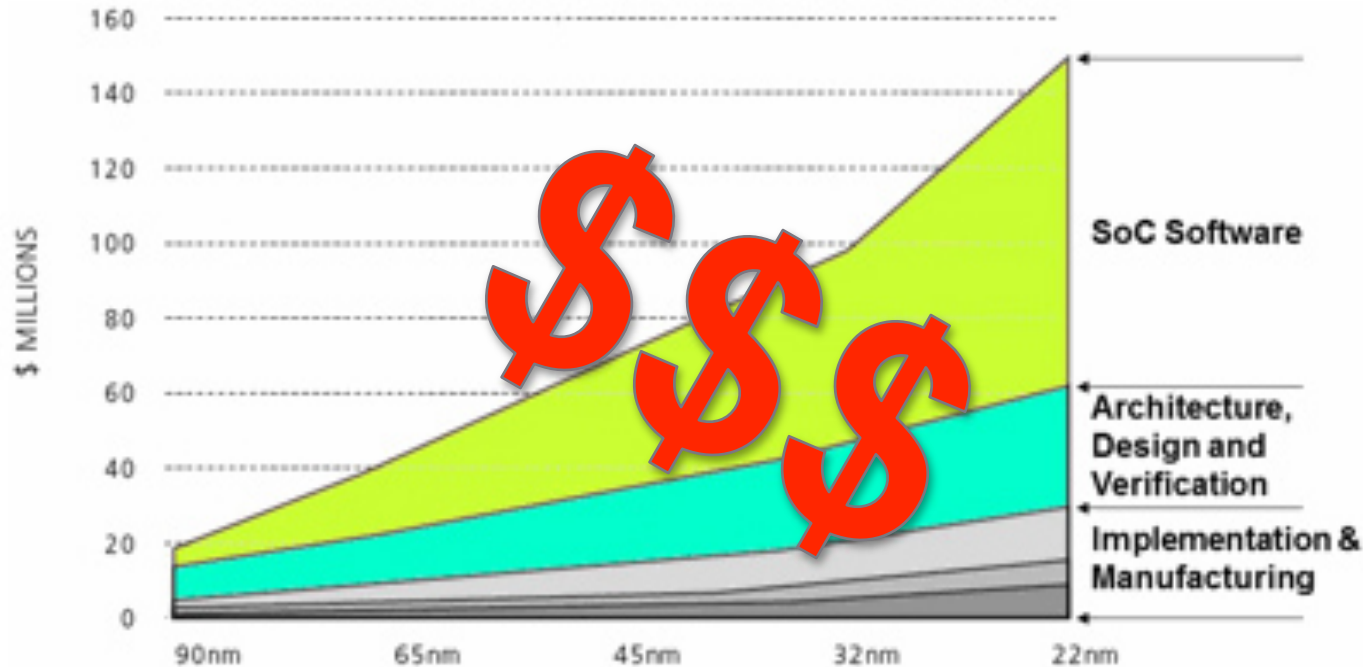


Q2. 3D V-NAND Can Solve ?

- At 20nm, about  $70e^-$  stored
- $V_t$  not ok for multilevel storage



# And the development cost is increasing



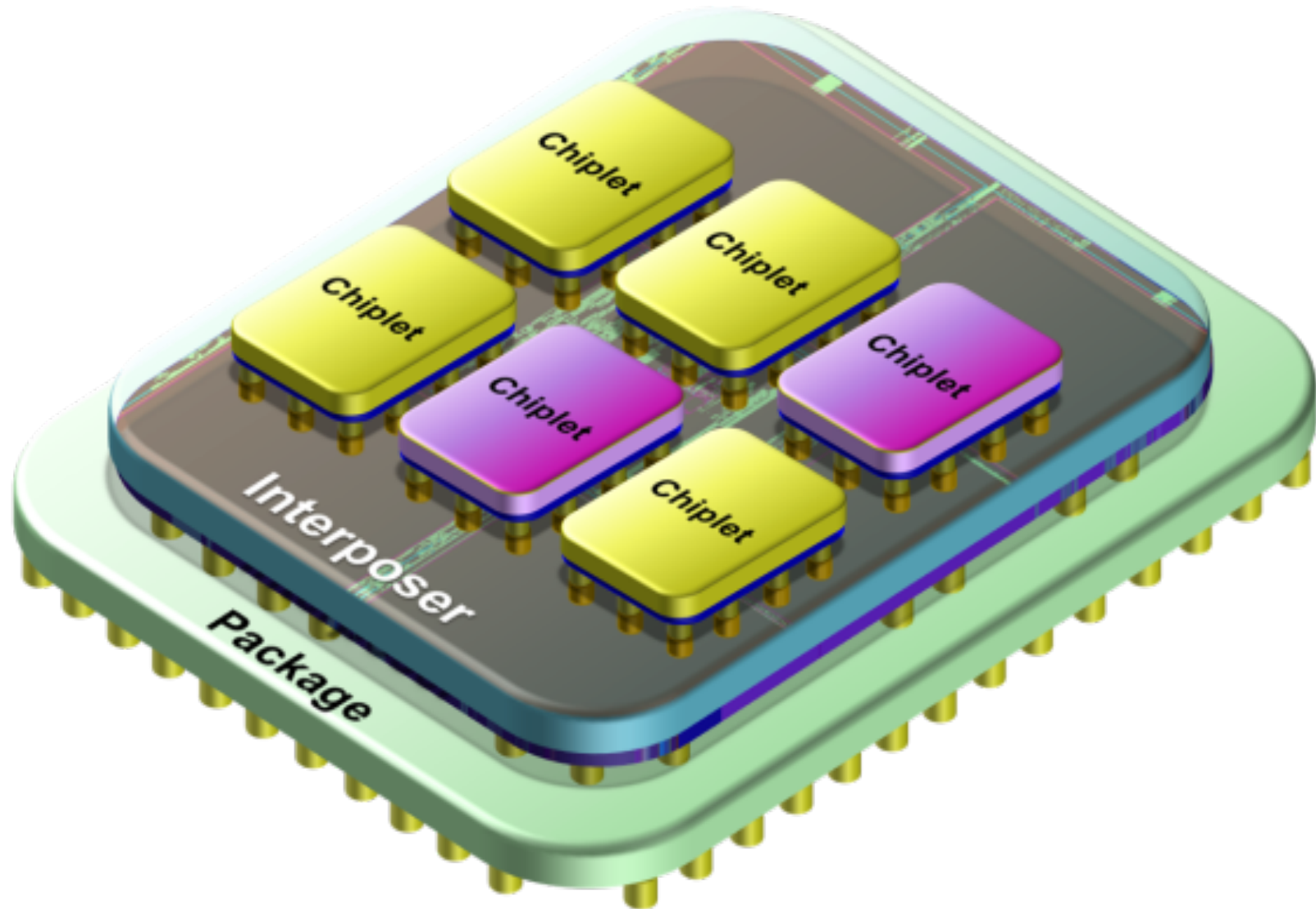
Source: International Business Strategies, Inc. (Los Gatos, CA)

**SoC Development Costs have Soared from \$20 Million at 90nm to Over \$100 Million at 32 nm**

Rock's law: cost of IC plant doubles every 4 years  
Reaching 10<sup>th</sup> of \$ Billions...

(Samsung will spend about \$15 billion to open a new semiconductor factory in South Korea by 2017).

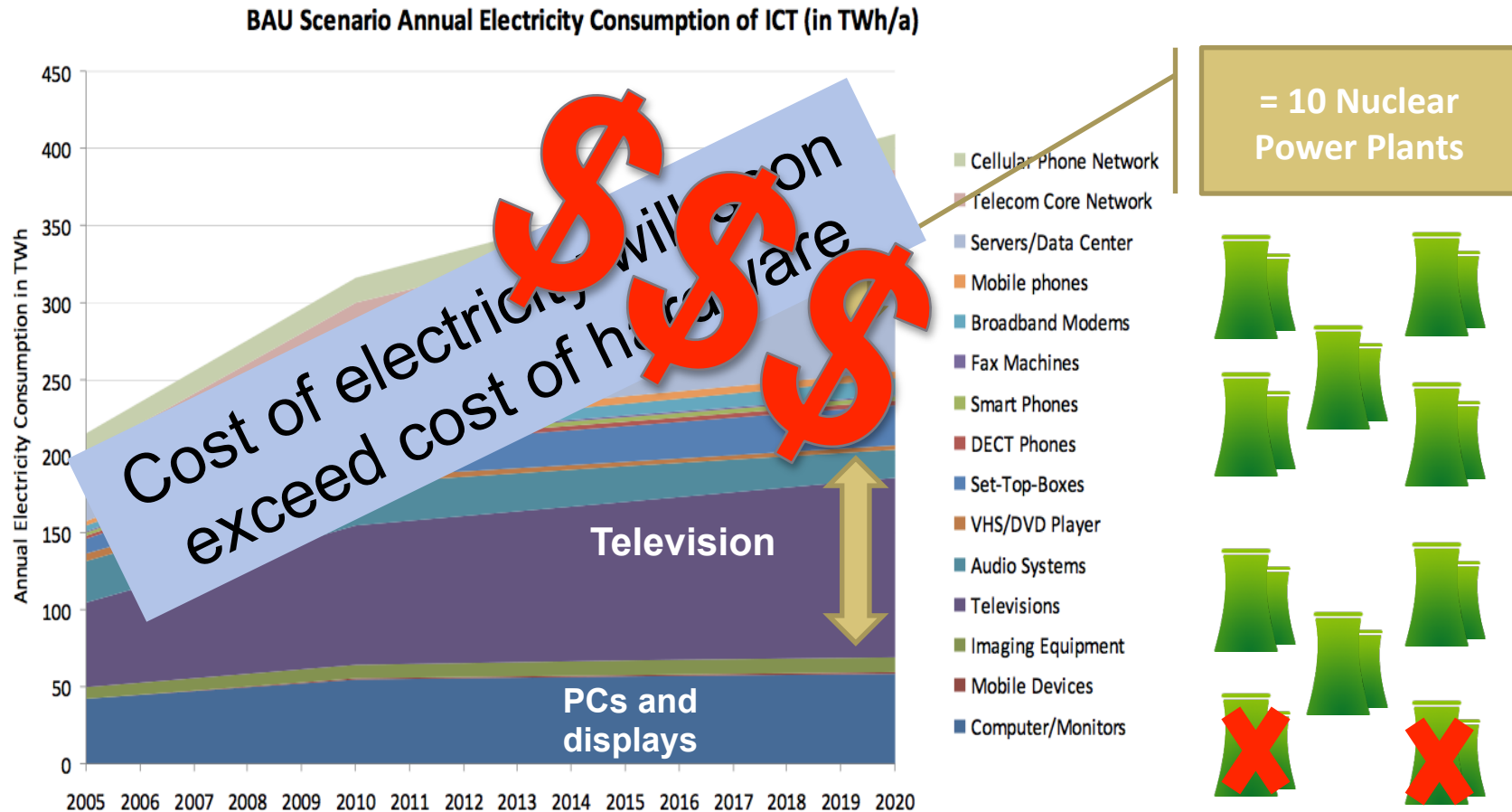
# Specialization with interposer





# Energy consumption of ICT

- Estimated consumption 410 TWh in 2020, 25% for servers



Source: European Commission DG INFSO, Impact of Information and Communication Technologies on Energy Efficiency, final report, 2008

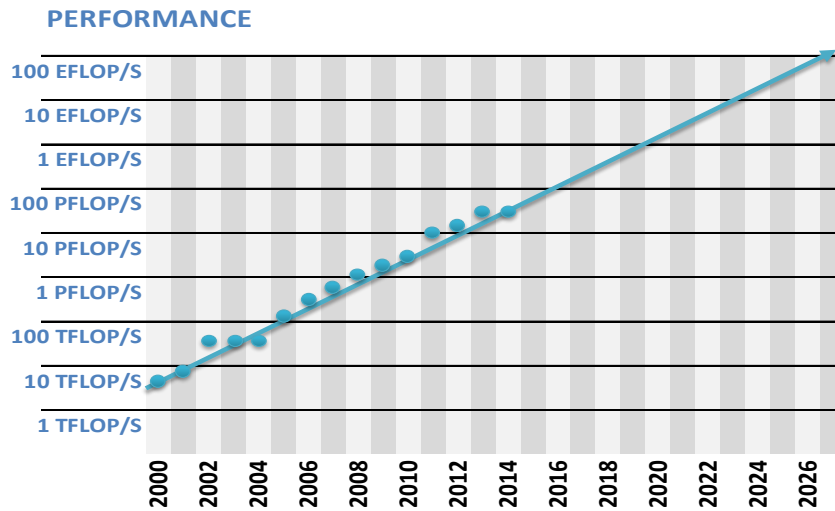
# The energy challenge for HPC

Power for CPU-only Exaflop Supercomputer   = Power for the Bay Area, CA (San Francisco + San Jose)

**HPC's Biggest Challenge: Power**

Source: Timothy Lanfear, « GPU computing and the future of HPC »

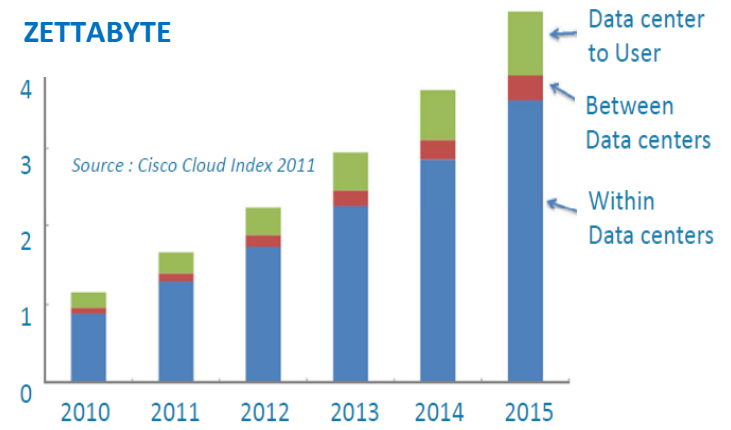
# Top Challenges for Computing



Projection from www.top500.org

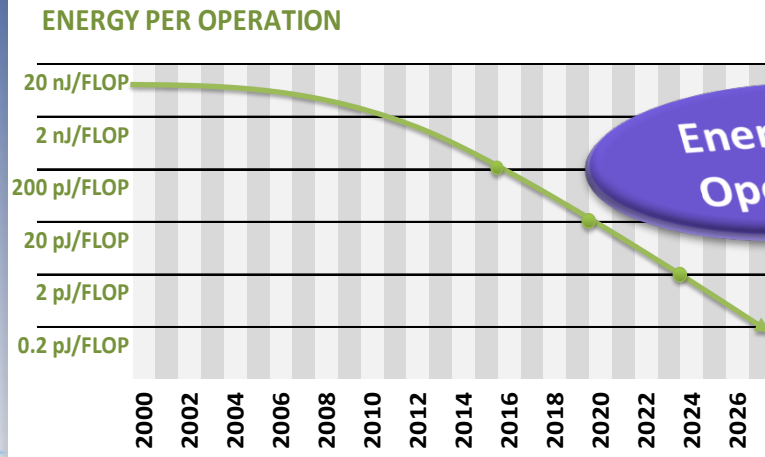
Exascale computing before 2020.

**Performance**



**Data Traffic**

Doubling every 2 years



Projection from www.top500.org

**Energy per Operation**

Divided by 4 every 2 years

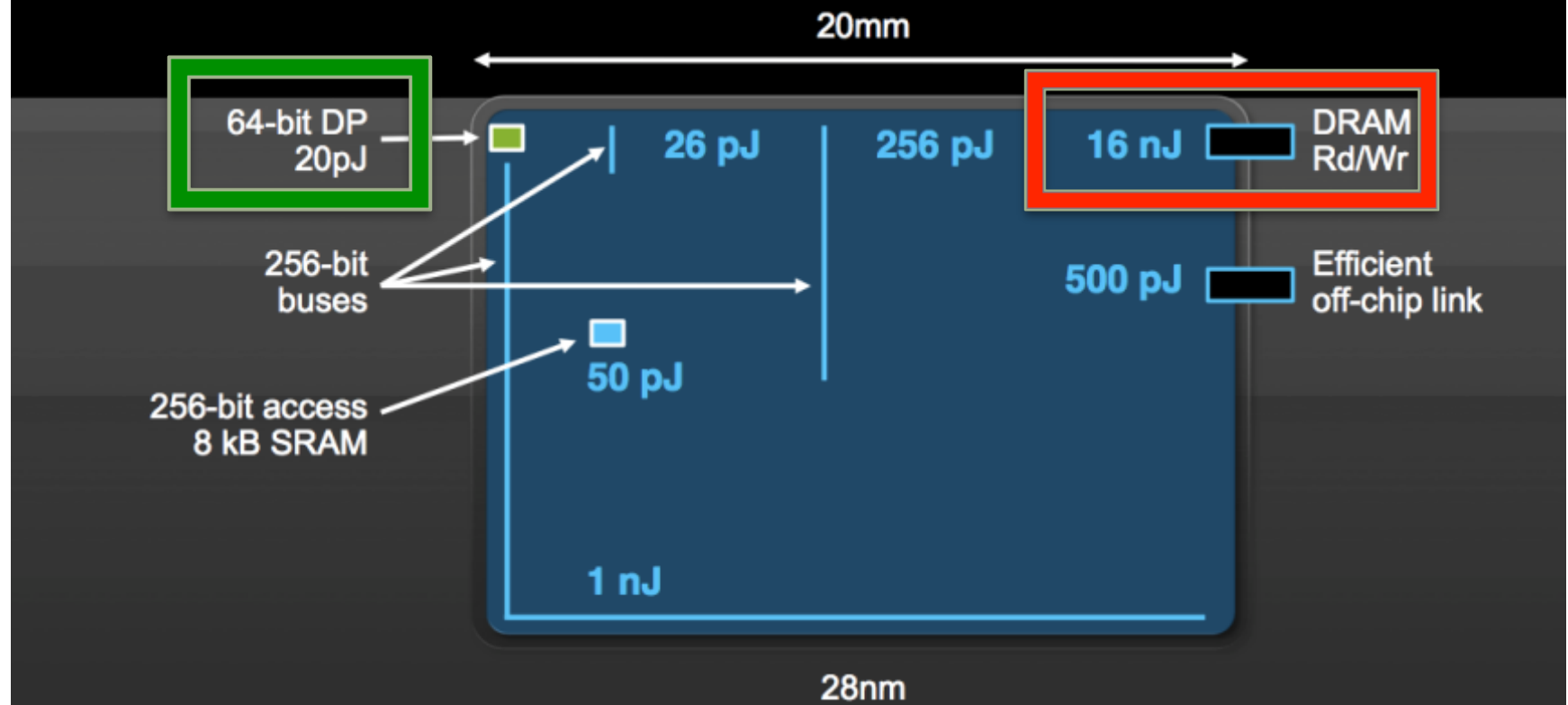
**Disruptive solutions are required**



# Cost of moving data

## The High Cost of Data Movement

Fetching operands costs more than computing on them



Source: Bill Dally, « To ExaScale and Beyond »

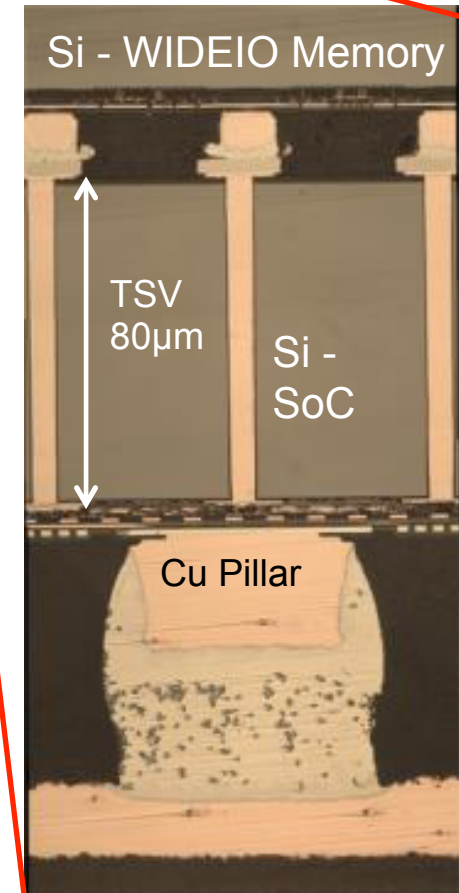
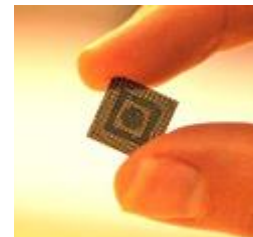
[www.nvidia.com/content/PDF/sc\\_2010/theater/Dally\\_SC10.pdf](http://www.nvidia.com/content/PDF/sc_2010/theater/Dally_SC10.pdf)

# Using the 3<sup>rd</sup> dimension: 3D stacking



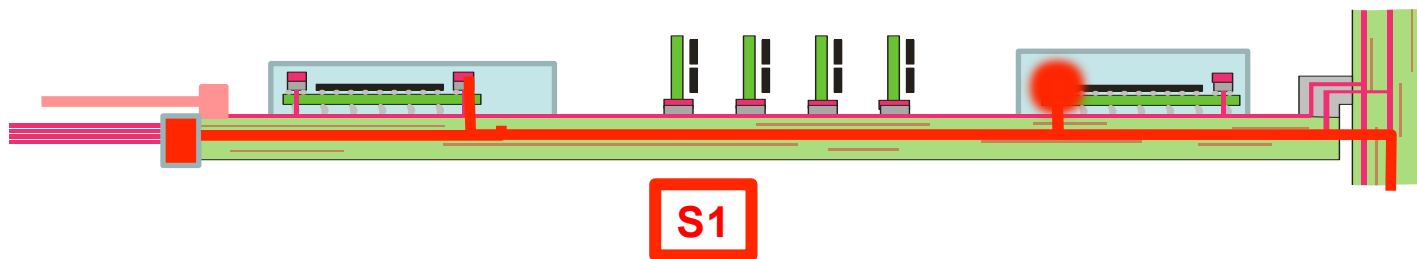
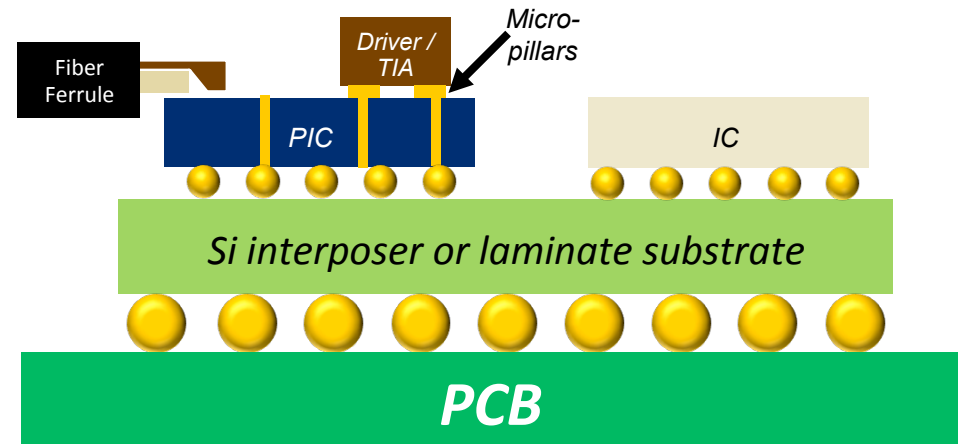
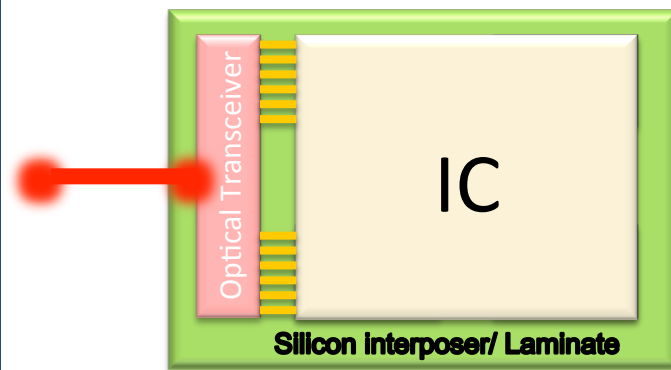
Exemple: WIDEIO memory stacked on top of a MPSoC in the same package

- Partnership between CEA-LETI, STEricsson, STMicroelectronics and Cadence
- High bandwidth: WideIO provides more than **34 Gbytes/s** (Currently: 17 GBytes/s)
- Low power: **4x power efficiency** compared to LPDDR2/3
- Compatible with FD-SOI
- FBGA Package
  - Size 12x12mm, Ball Pitch 0.4mm, 1.2 mm thickness





# Off-chip photonics



Off board: AOC, optical modules

Now

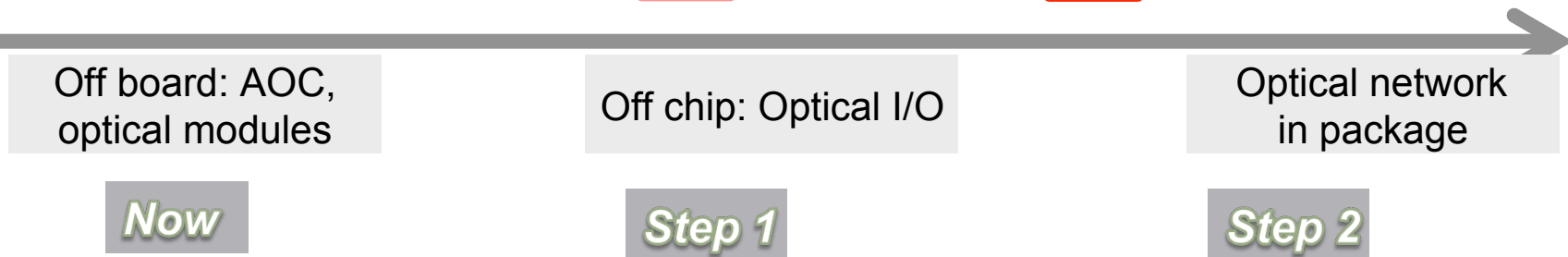
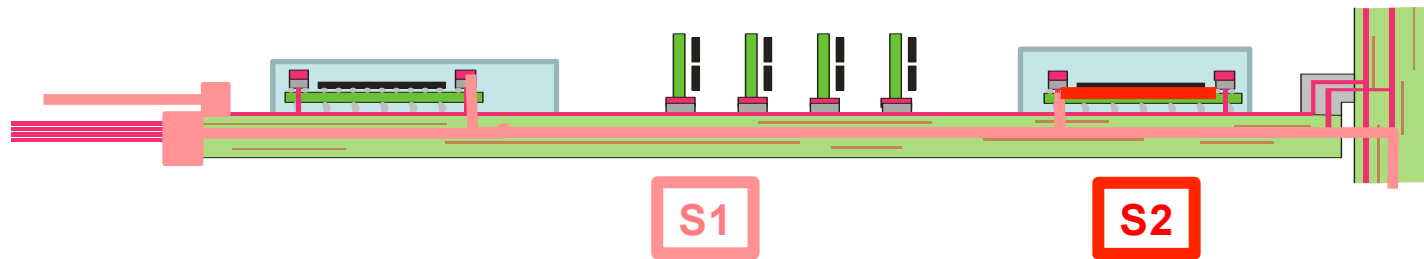
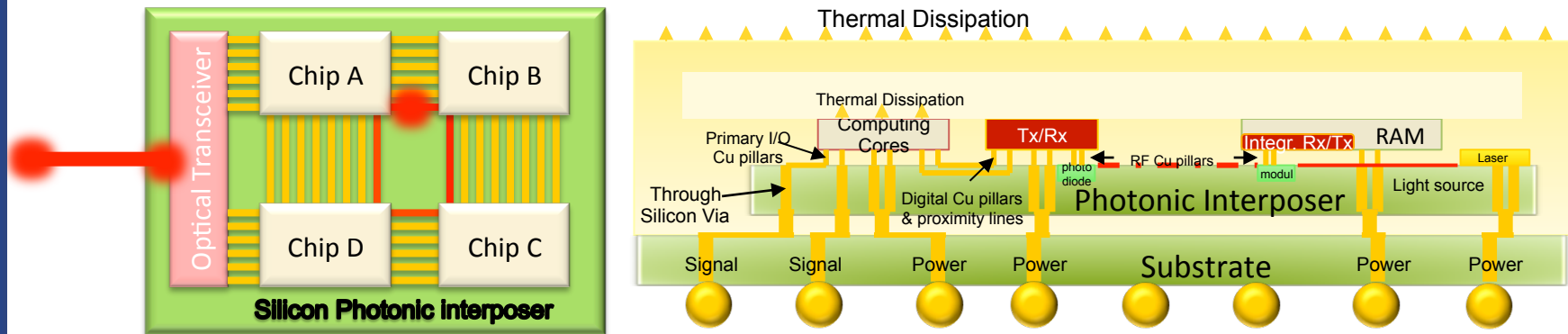
Off chip: Optical I/O

Step 1

Time



# In-package photonics



# Where are we ?

We can still put more transistors per mm<sup>2</sup> for the coming few years

- But energy is a key limiting factor
  - New technologies (FinFet, FDSOI)
  - 3D stacking
  - More efficient architectures, coprocessors
- SRAM, DRAM didn't scale anymore
- Flash is running out of electrons
- Kryder's law for Hard Disk Drives (forecast 40% increase density per year, reality 15%)
- Non-volatile memories are promising
  - But which technology, at which density and reliability?
  - Still time required for replacing completely "old" storage technologies



# The future will be non volatile memories

But still in development and which technology will be the winner?

	FeRAM	RRAM	Magnetic field write MRAM	PRAM	STT MRAM
Non-volatile	Y	Y	Y	Y	Y
Memory cell factor ( $F^2$ )	16-32	4-6	16-32	5-8	5-7
Read time (ns)	20-50	10-20	3-20	5-20	3-15
Write/erase time (ns)	50	20	10-20	>30	3-15
Number of rewrites	$10^{12}$	$10^5$	$10^{15}$ min	$10^{12}$	$10^{15}$ min
Power consumption at write	Low	Low	Somewhat high	Low	Low
Required input voltage (V)	2-3	1.2	3	1.5-3	1.5



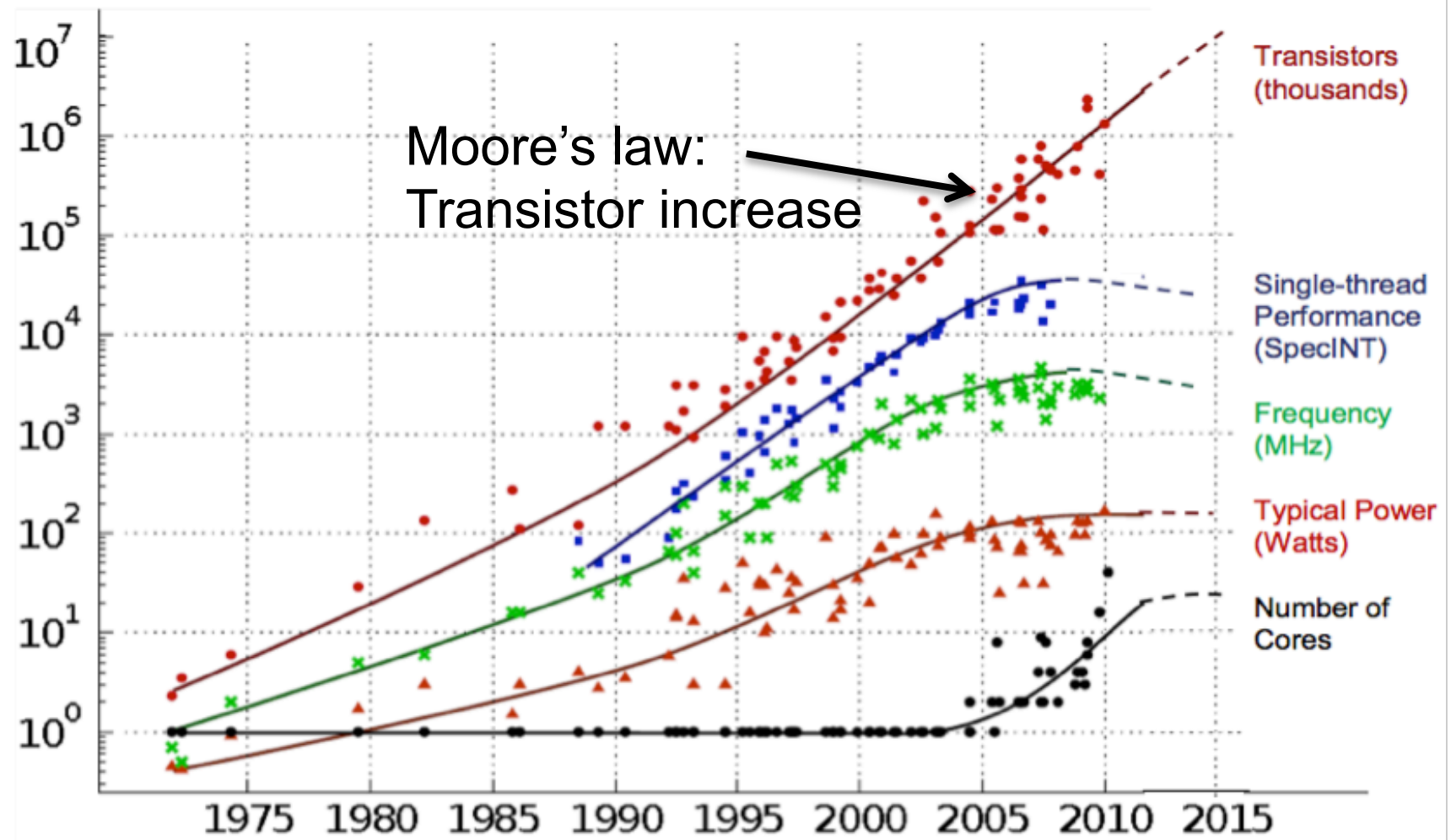
# Status of “laws”



- “Laws” not true anymore:
  - **Dennard’s law** (1974): As transistors get smaller their power density stays constant, so that the power use stays in proportion with area
  - **Kryder’s law** (2005): The magnetic disk area storage density is increasing 40% per year (2009)
- “Laws” still on-going:
  - **Moore’s law** (1965): The density of transistors is doubling every year (1965) or every two years (1975)
  - **Rock’ law**: The cost of a semiconductor chip fabrication plant doubles every four years.
  - **Amdahl’s law** (1967): the speedup of a program using multiple processors in parallel computing is limited by the time needed for the sequential fraction of the program.
  - **Gustafson’ law** (1988): Computations involving arbitrarily large data sets can be efficiently parallelized.



# Moore's law still on-going



Source from C Moore, « Data Processing in ExaScale-Class Computer Systems », Salishan, April 2011

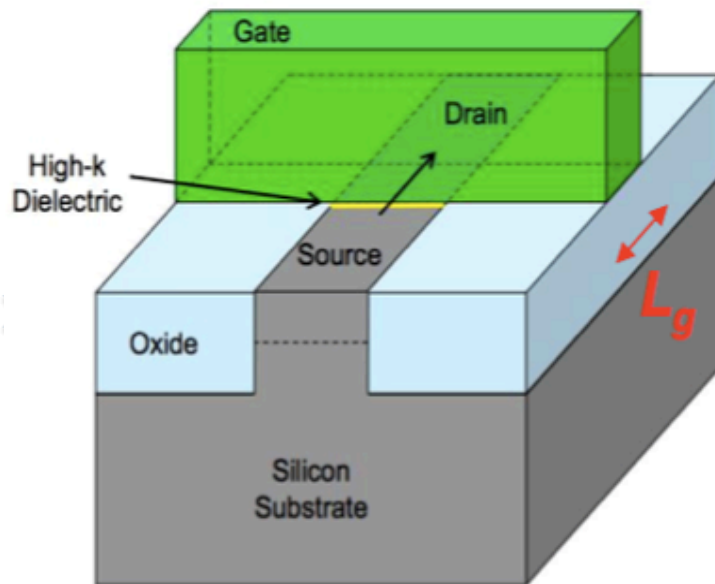


# Moore's law still on going...

## Advanced CMOS roadmap

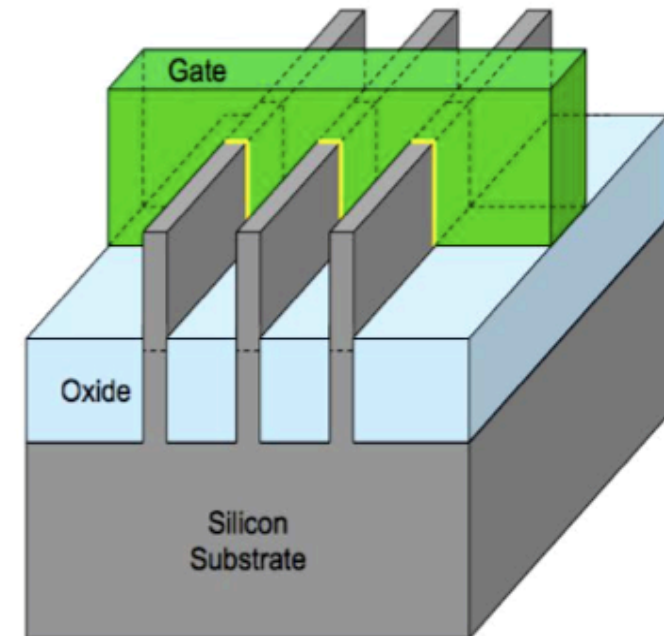
Evolutionary scaling: technology driven performance improvement

### Transistor 2D



Current control difficult when  $L_g < 20\text{nm}$

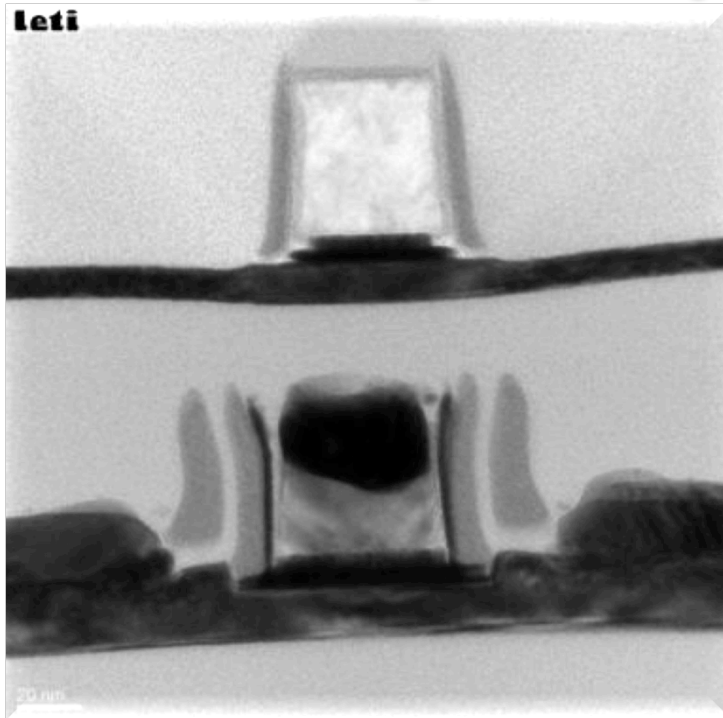
### Transistor « Tri-gate » 3D



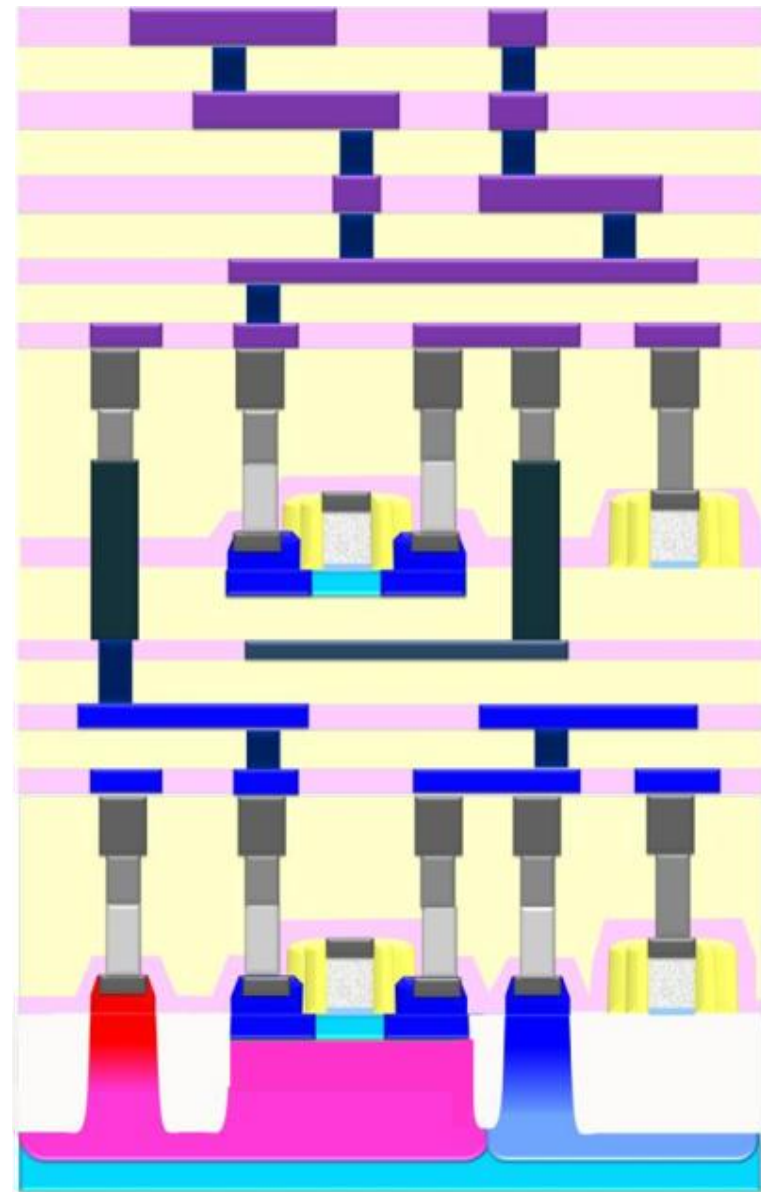
Better control (2 sides + multiple gates)



# M3D principle



CMOS/CMOS: 14nm vs 2D:  
Area gain=55%  
Perf gain = 23%  
Power gain = 12%



# Technology Roadmap



## Litho

Water Drain, Lens, Water Source, Si Wafer, Wafer Stage, Source, Mask, Lightening, Co-optimize, OPC, Double Patterning(LELE), Triple Patterning, DSA, SPT Process, QSPT Process, EUV

193nm Imm.ArF    Source Mask Optimization    Double Patterning(SADP)    SAQP

## Front End

歪みシリコン技術, 内部応力の向き, PMOS, HKMG GL, NMOS SiC:P, SiGe S/D, channel, SiGe S/D, 6nm SiGe Ch, gate, SiC RS, InGaAs, Horizontal, MoS2, SiO2, Si, 2D Material, Graphene FET, FDSOI, Fin FET, FDSOI(Strain), Si, B0x, SiC:P, Weff boost Fin FET, Ge channel, III-V channel, Vertical, Graphene Wiring

Stress Engineering    SiGe P-Ch.    Fin FET    FDSOI(Strain)    Ge channel    III-V channel    Vertical

## Back End

Ultra low-k, K=2.4~2.8, Air Gap, 14 nm Process, 52 nm pitch, CNT Via

Courtesy: Yuzo Fukuzaki – JAPAN PIDS



# Together...

## Electrons for compute

Electrons like to interact; easily moved; interaction needed for compute

## + Ions for storage

Ions like to interact; stay put; good for storage

## + Photons to communicate

Photons don't like to interact or stay put; good for long-distances

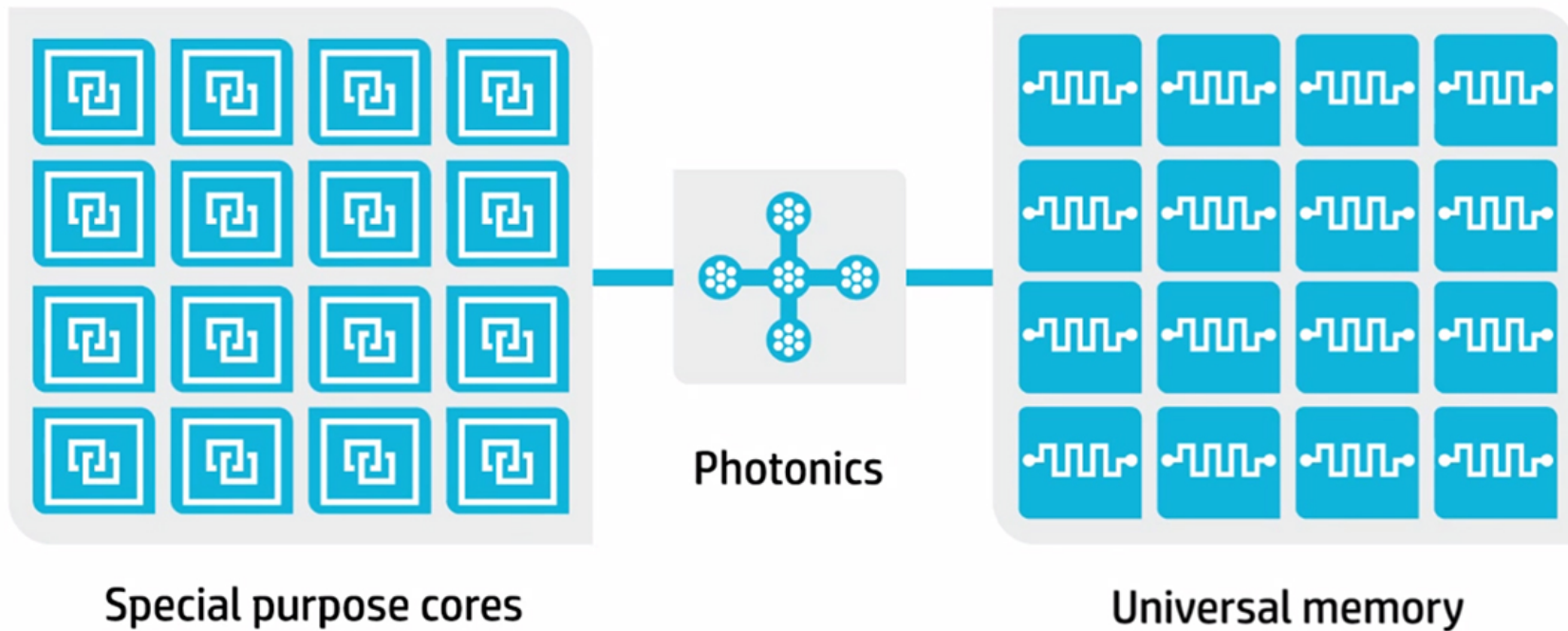
See the presentation on “The Machine” from HP:  
<https://www.youtube.com/watch?v=JzbMSR9vA-c>

Courtesy: Jouppi2011

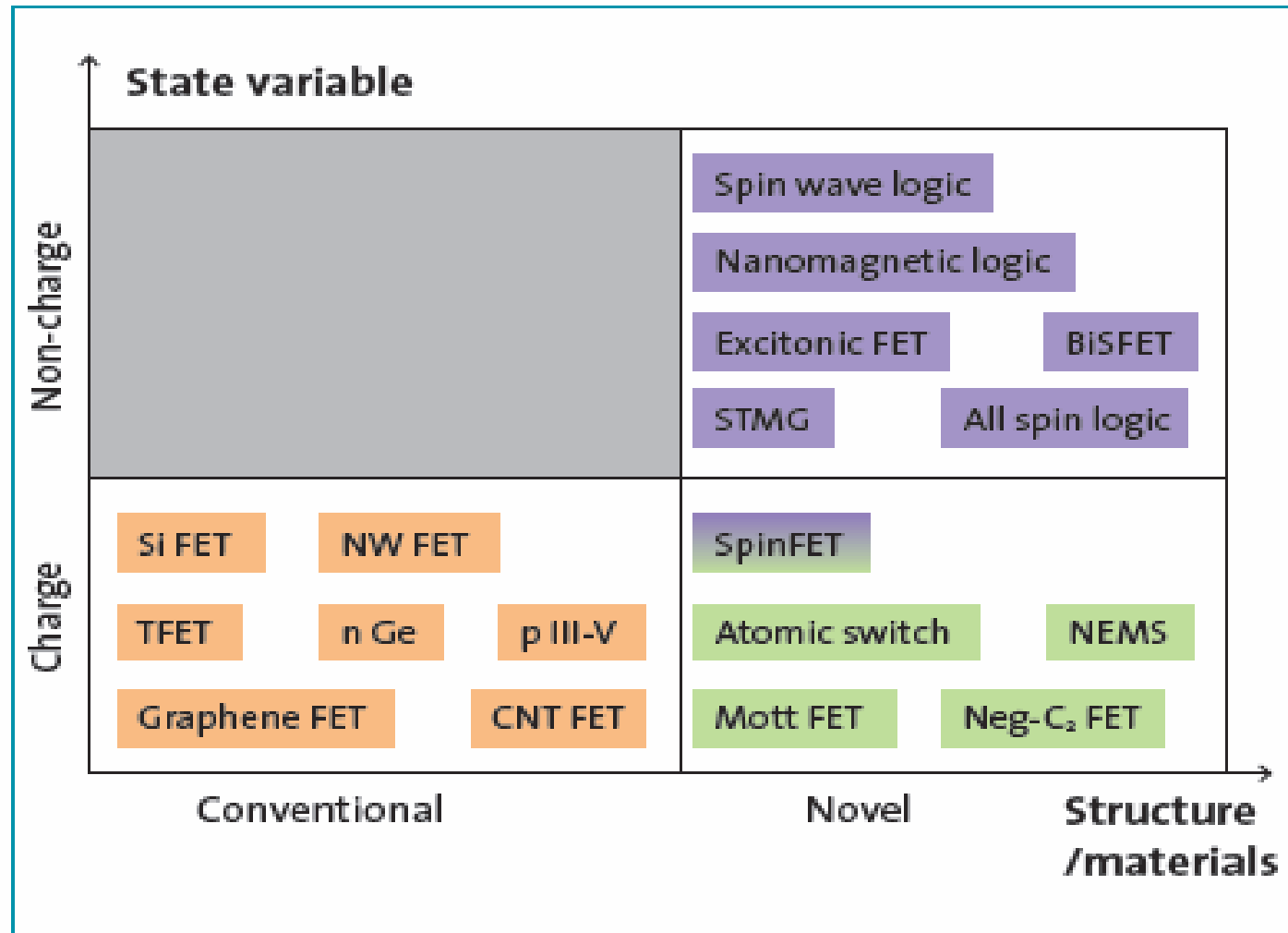


Source: P. Ranganathan, “Saving the world together, one server at a time...” ACACES 2011

# HP and “The Machine” announced in June 2014



# New emerging technologies for computing

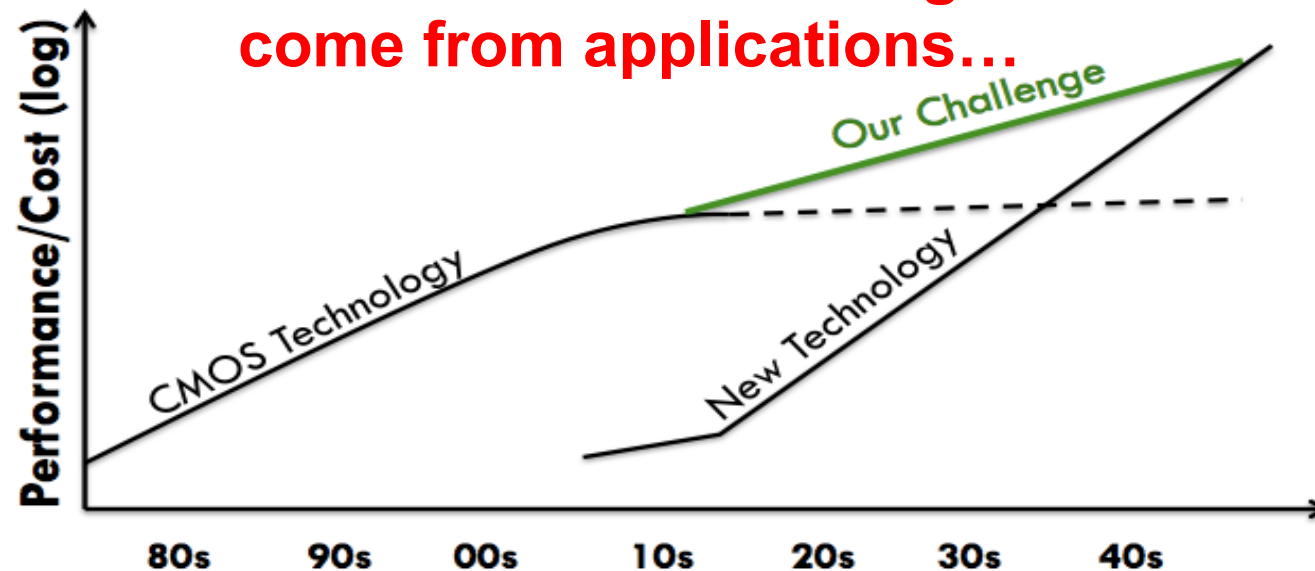


*Taxonomy for emerging information processing devices (from [ITRS2013]).*

# We are entering into a transition period...

## Scaling without Technology Help

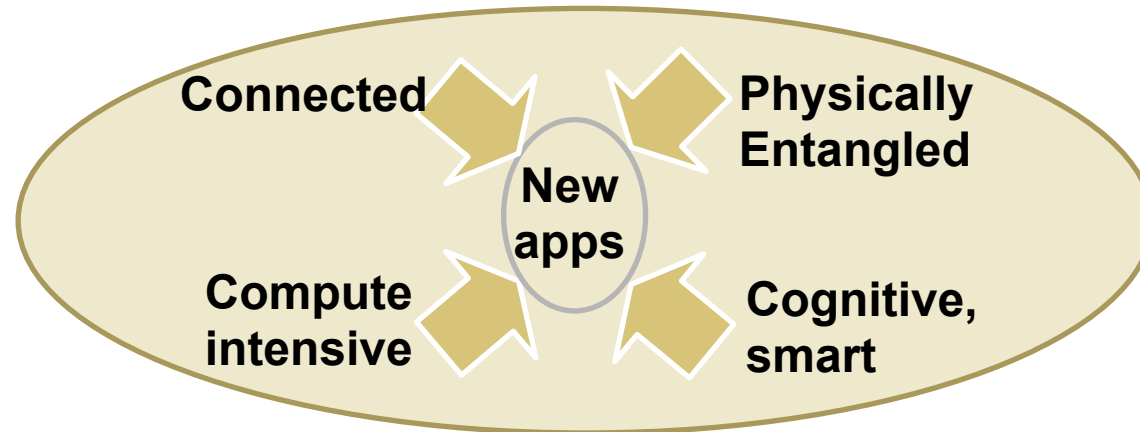
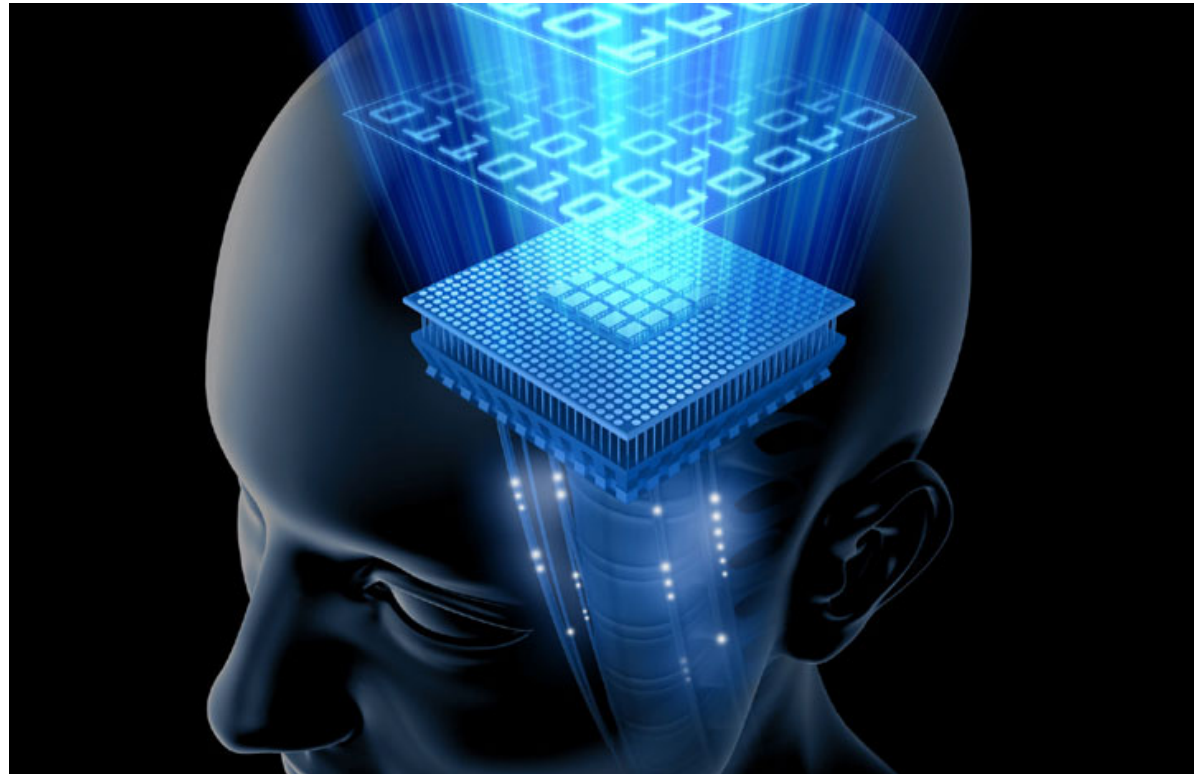
**But even more challenges  
come from applications...**



[Hill & Kozyrakis'12]

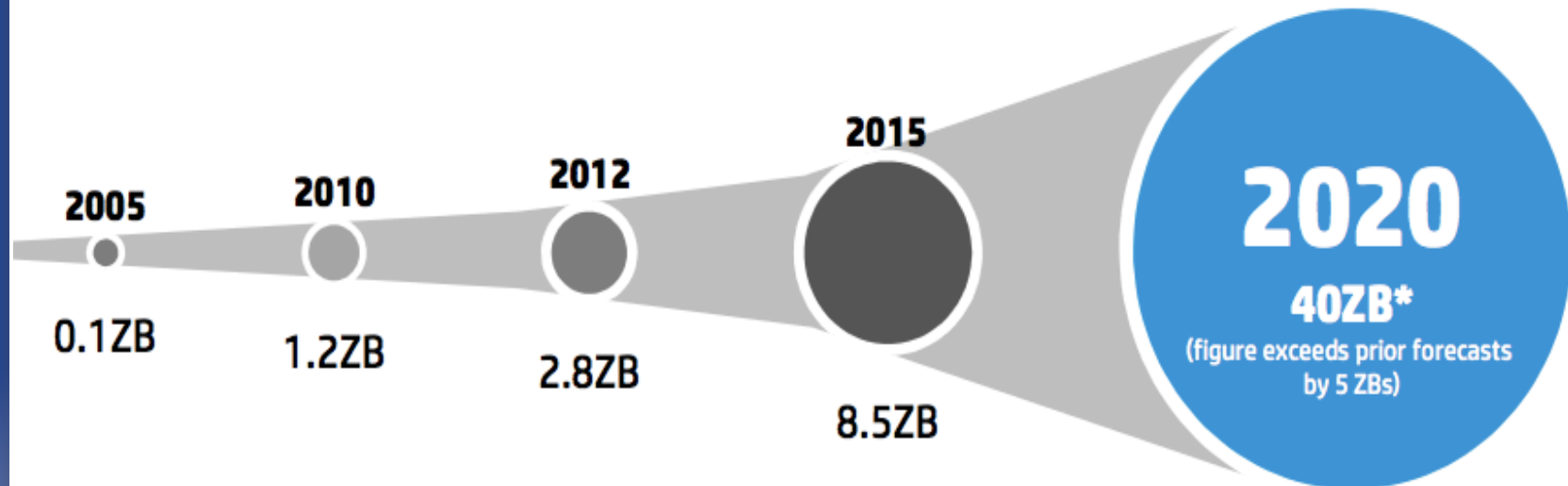


# New apps will be...





# The data deluge challenge

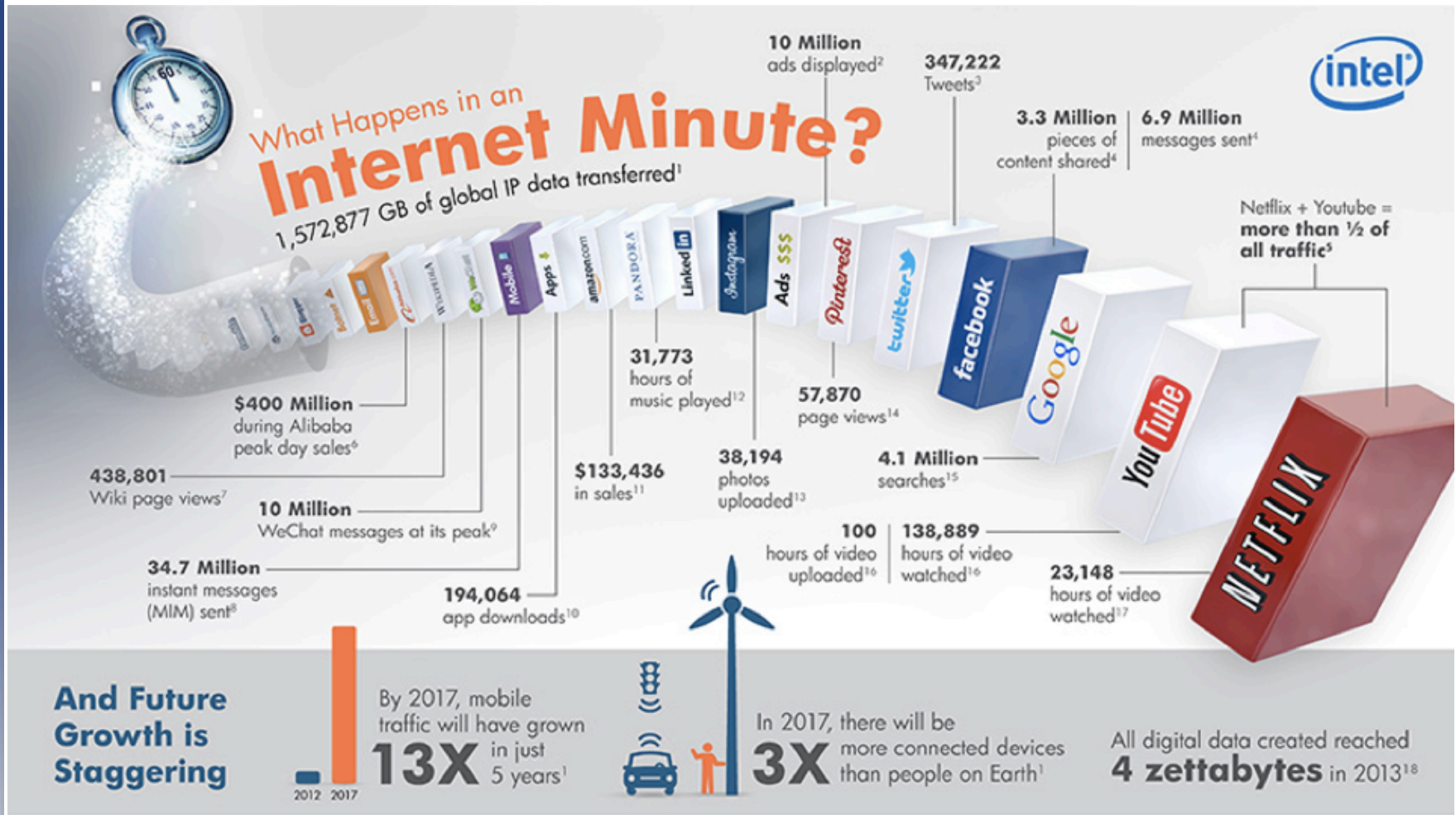


**1 ZB =  $10^{21}$  bytes**

40 ZB is equal to 57 times the amount of all the grains of sand on all the beaches on earth.

Source: Paolo Faraboschi, HP, and IDC

# Data creation every minute in 2014



# IoT: the Internet of **Threats**

- Today security / privacy issues make the newspaper headlines

**Massive adoption of IoT by citizens relies on confidence in terms of security and privacy**

**CBSN**  
By CHENDA NGAK / C  
**Can you?**

**The A**  
Data Centre Software Networks

**IEEE SPECTRUM**  
Follow on: f t in +  
Podcasts | Biomedical | Device  
**Hacking Pac...**

**CNN**  
Sign up | Log in  
SEARCH  
POWERED BY Google  
Home Video World U.S. Africa Asia Europe Latin America Middle East Business  
**Connected TVs, fridge help launch global cyberattack**  
By Brandon Griggs, CNN  
January 17, 2014 -- Updated 2252 GMT (0652 HKT) | Filed under: Gaming and Gadgets

**SECURITY**  
**Pol...**

A Polish teenager allegedly turned the tram's personal train set, triggering chaos and derailed people were injured in one of the incidents.

The 14-year-old modified a TV remote control's points, *The Telegraph* reports. Local police said depots to gather information needed to build the he modified track setting for a prank.

1 Comment / 36 Shares / 6 Tweets / Stumble / @ Em

**LAS VEGAS** Closing the curtains in your living room may not potential hackers. At a demonstration Friday in Las Vegas, researchers showed an audience of children at Defcon Kids how a Samsung Smart TV can be hacked.

A smart refrigerator on display at the International Consumer Electronics Show (CES) last week in Las Vegas.

# Misuse of information technology might destroy our privacy



Snowden effect



Heartbleed bug - OpenSSL

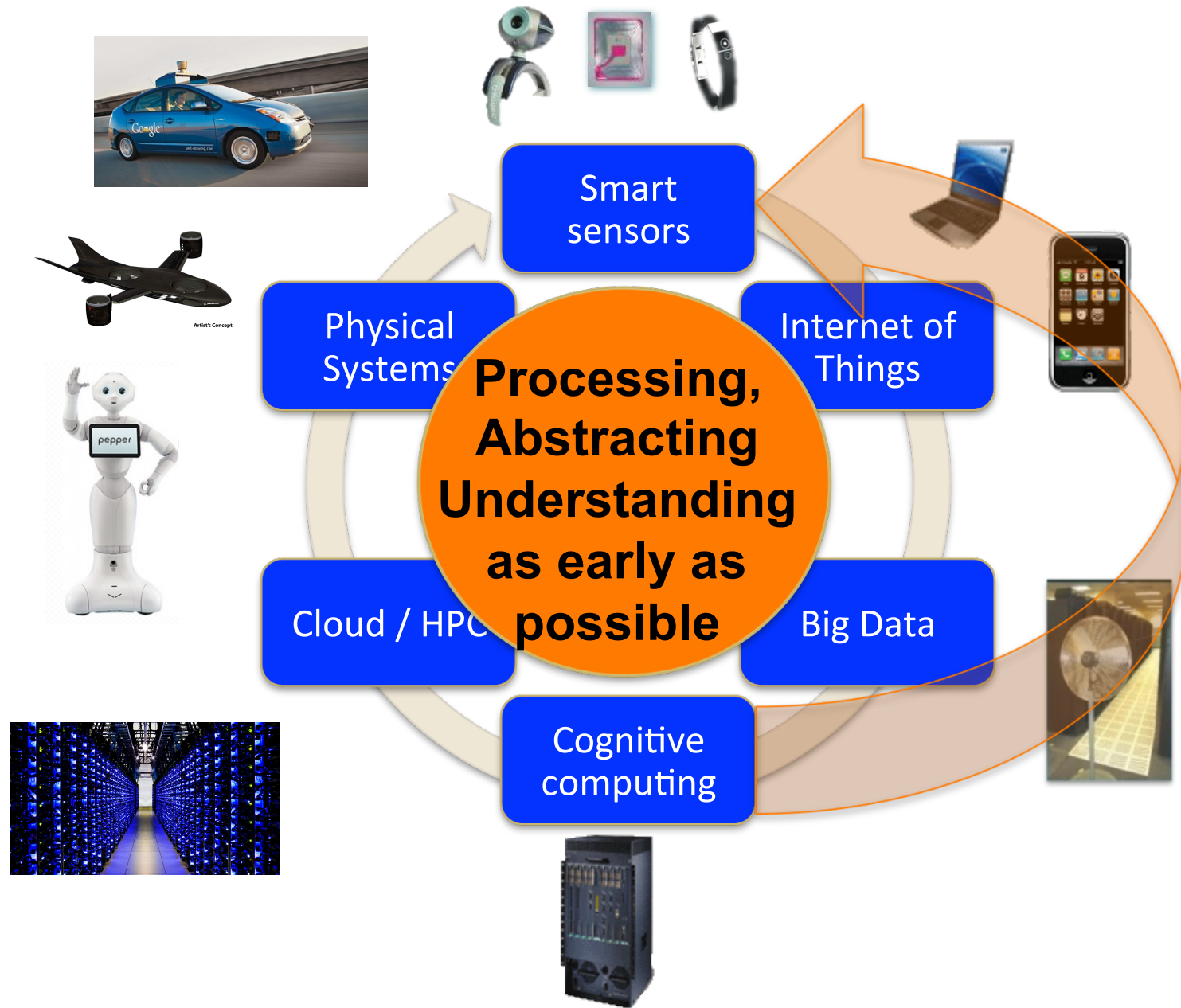


Internet of things

- Consumers are willing to give away private information for free services



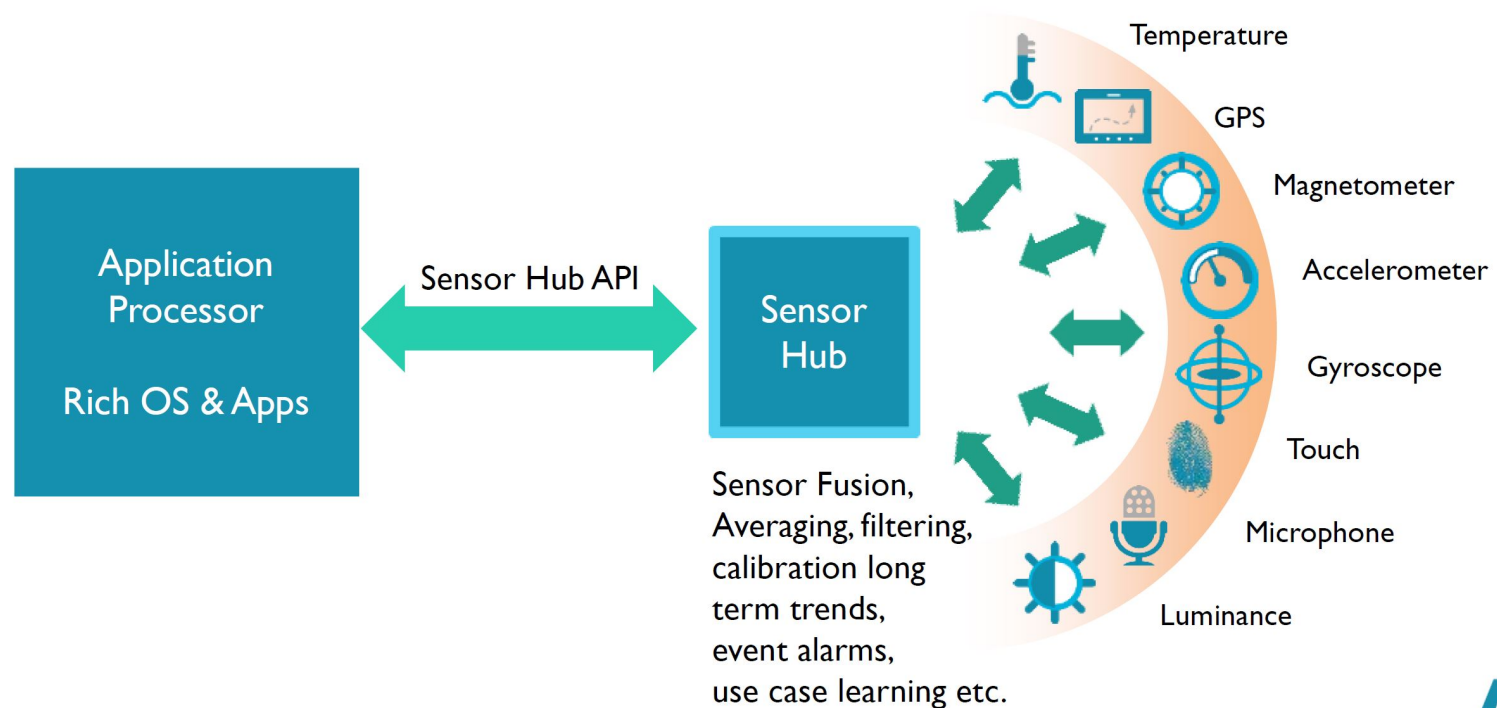
# Global integration of communication, computation and reaction



# Example of architecture for end-nodes

## 'Always-on, Always-Aware' Architecture key for Sensors

Sensor Hub brings contextual awareness tracking even when the Apps Processor is in standby



ARM®



# Computing becomes increasingly cognitive

« Tomorrow's cognitive systems will be fundamentally different from the machines that preceded them. While traditional computers must be programmed by humans to perform specific tasks, **cognitive systems will learn from their interactions with data and humans** and be able to, in a sense, **program themselves to perform new tasks.**

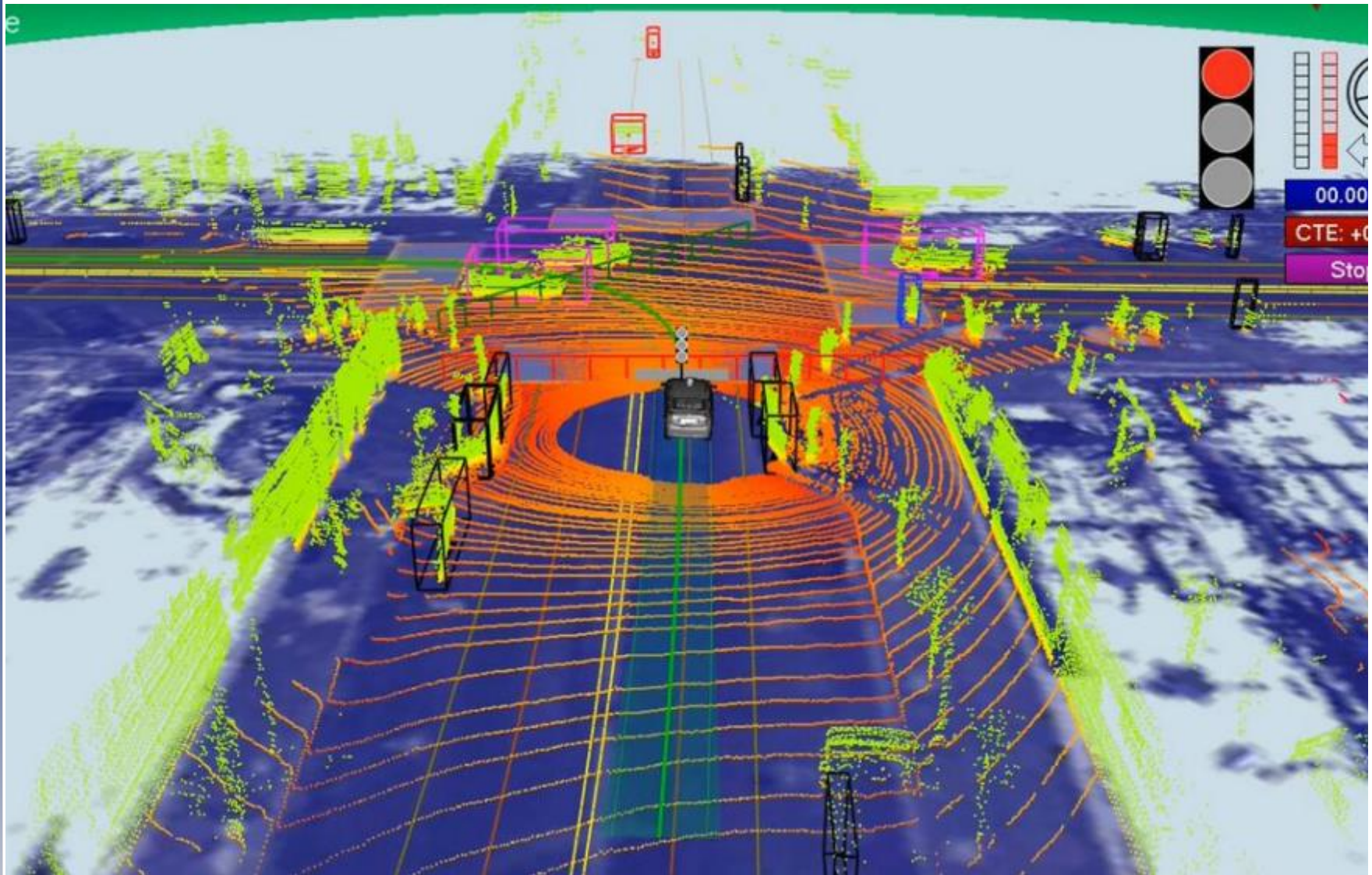
Traditional computers are designed to calculate rapidly; **cognitive systems will be designed to draw inferences from data and pursue the objectives they were given.** <...>

Cognitive systems will augment our hearing, sight, taste, smell, and touch. In the programmable-computing era, people have to adapt to the way computers work. In the cognitive era, **computers will adapt to people.**

They'll interact with us in ways that are natural to us. »

From Kelly III, John E. and Hamm, Steve. « Smart Machines: IBM's Watson and the Era of Cognitive Computing. »

# Integrated in the physical reality (CPS)



Ramin Rahimian for The New York Times

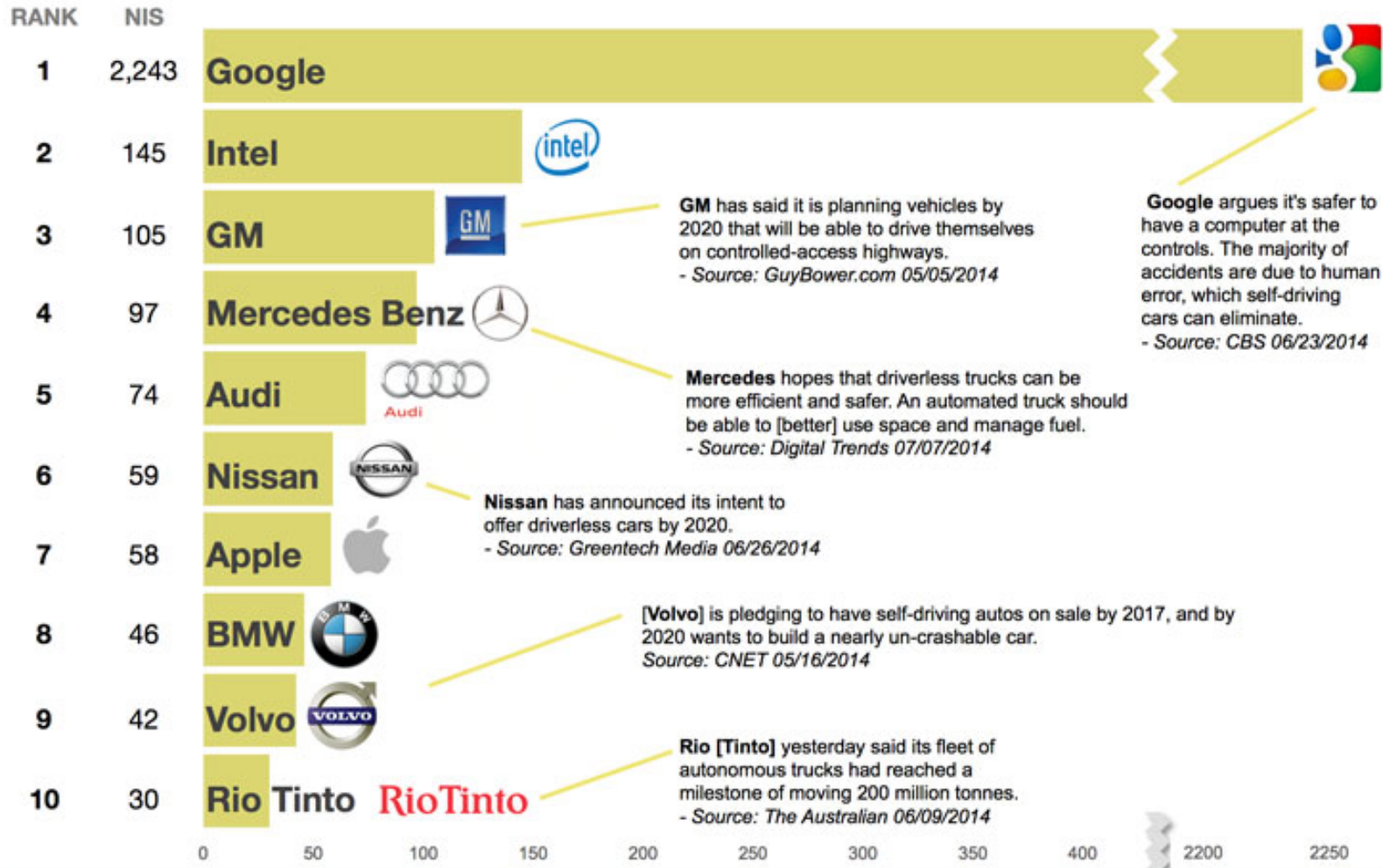
Dmitri Dolgov, a Google engineer, in a self-driving car parked in Silicon Valley after a road test.



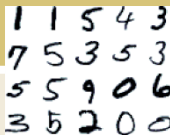







# Autonomous vehicle: will it induce a change in industry?

## The 10 Most Influential Autonomous Cars Companies



# Deep Neural Networks: state-of-the-art in image recognition

Database	# Images	# Classes	Best score
MNSIT <i>Handwritten digits</i> 	60,000 + 10,000	10	99.79% [3]
GTSRB <i>Traffic sign</i> 	~ 50,000	43	99.46% [4]
CIFAR-10 <i>airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck</i> 	50,000 + 10,000	10	91.2% [5]
Caltech-101 	~ 50,000	101	86.5% [6]
ImageNet 	~ 1,000,000	1,000	Top-5 83% [1]
DeepFace 	~ 4,000,000	4,000	97.25% [2]

INCREASING COMPLEXITY

- State-of-the-art are Deep Neural Networks *every time*

# Neural Network and image applications

Convolutional Neural Network : state of the art performance on several databases for image classification:

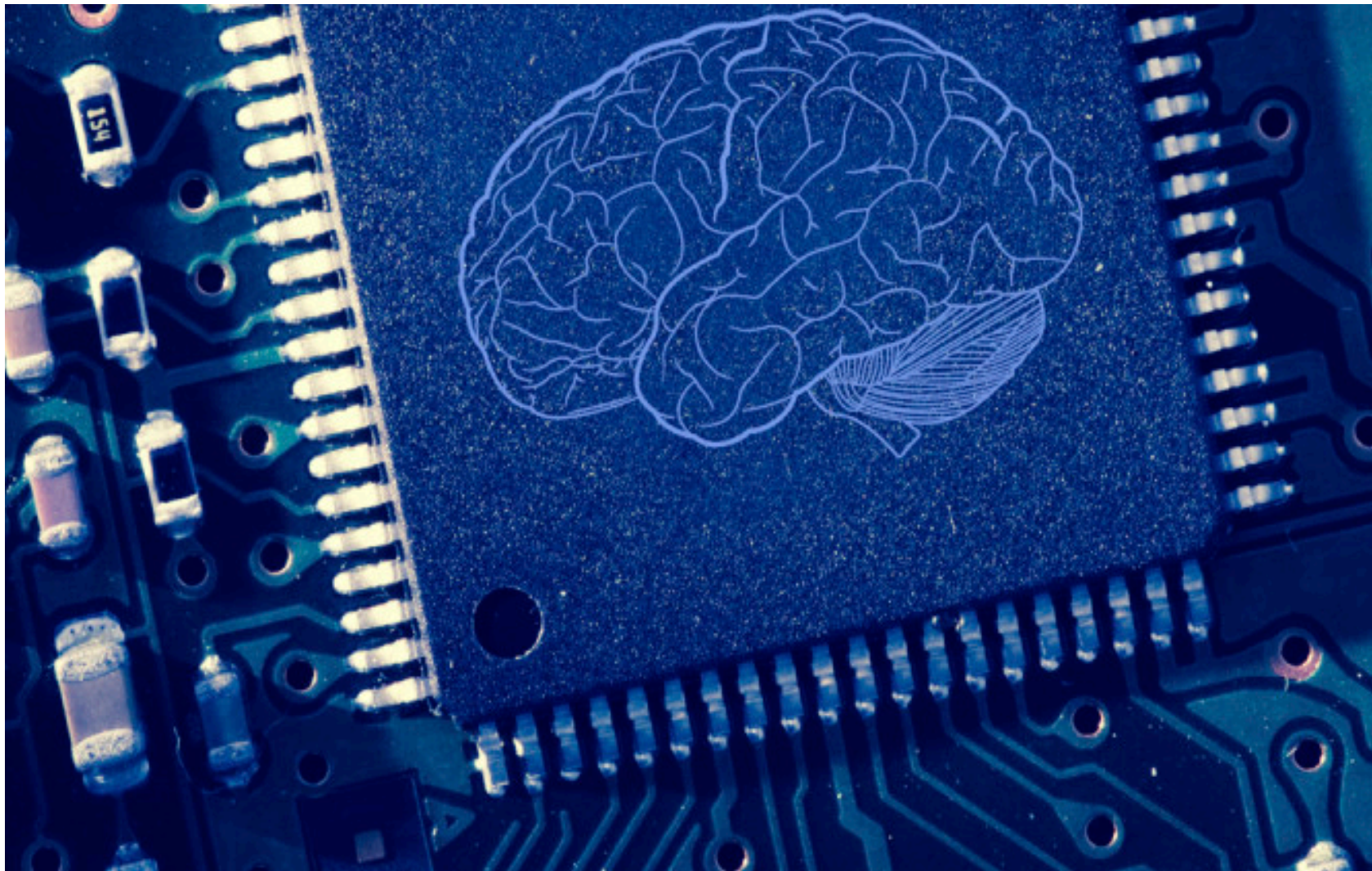
- **MNIST**: 99,77% good results (Comité CNN)
- **GTSRB**: 13/14 best results from CNN
- **CIFAR 10** : 89% good results
- **Image Net (ILSVRC)**

#	Team	Method	Accuracy
	<b>sermanet</b>	<b>EBLearn 2LConvNet ms 108 feats + 100-feats CF classifier + No color</b>	<b>99.17%</b>
197	IDSIA	cnn_hog3	98.98%
196	IDSIA	cnn_cnn_hog3	98.98%
<b>178</b>	<b>sermanet</b>	<b>EBLearn 2LConvNet ms 108 feats</b>	<b>98.97%</b>
195	IDSIA	cnn_cnn_hog3_haar	98.97%
<b>187</b>	<b>sermanet</b>	<b>EBLearn 2LConvNet ms 108 + val</b>	<b>98.89%</b>
199	INI-RTCV	Human performance	98.81%
170	IDSIA	CNN(IMG)_MLP(HOG3)	98.79%



Source: Yann LeCun ICML 2012

# Therefore, re-birth of Neural Networks



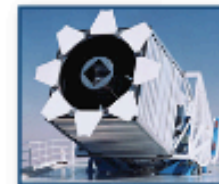
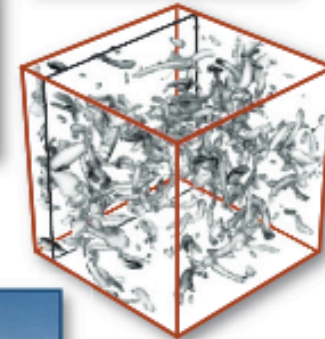
# The next step of research

## Science Paradigms

- Thousand years ago:  
science was **empirical**  
*describing natural phenomena*
- Last few hundred years:  
**theoretical** branch  
*using models, generalizations*
- Last few decades:  
a **computational** branch  
*simulating complex phenomena*
- Today: **data exploration** (eScience)  
*unify theory, experiment, and simulation*
  - Data captured by instruments  
or generated by simulator
  - Processed by software
- Tomorrow: (**Cognitive science?**)
  - Computers analyse databases/files  
and infer new discoveries.

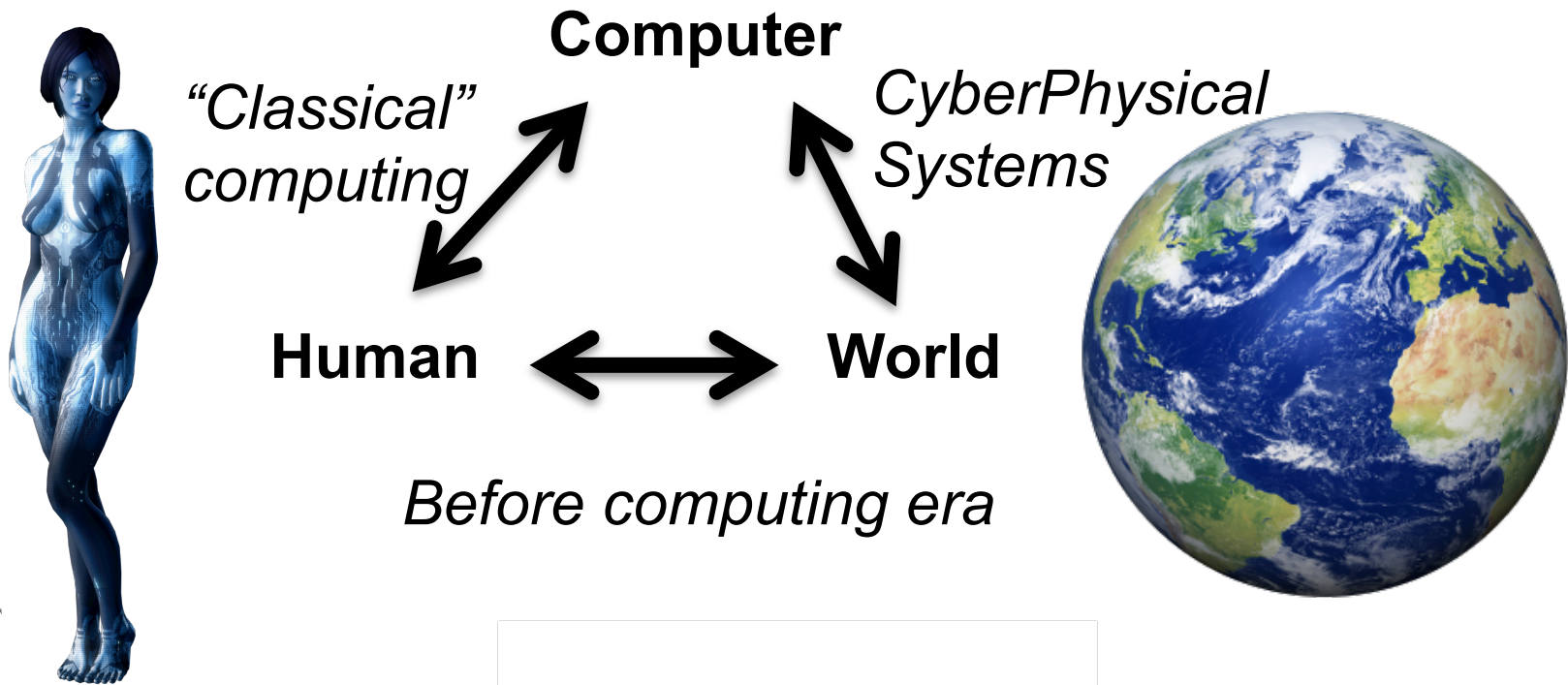


$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$





# Global integration of communication, computation and reaction





# Global integration of communication, computation and reaction

*Applications are delocalized, distributed on collaborating devices*



*Machine to Machine Interactions*



**Computer**

*Constraints of the real world e.g. time, ...*

**Human**

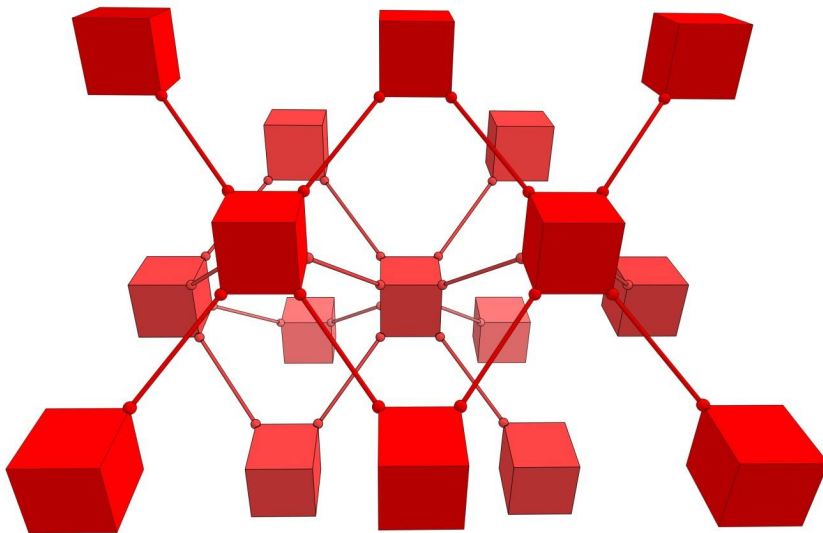


**World**



# Managing complexity....

*“Nontrivial software written with threads, semaphore, and mutexes is **incomprehensible** by humans”*



Edward A. Lee

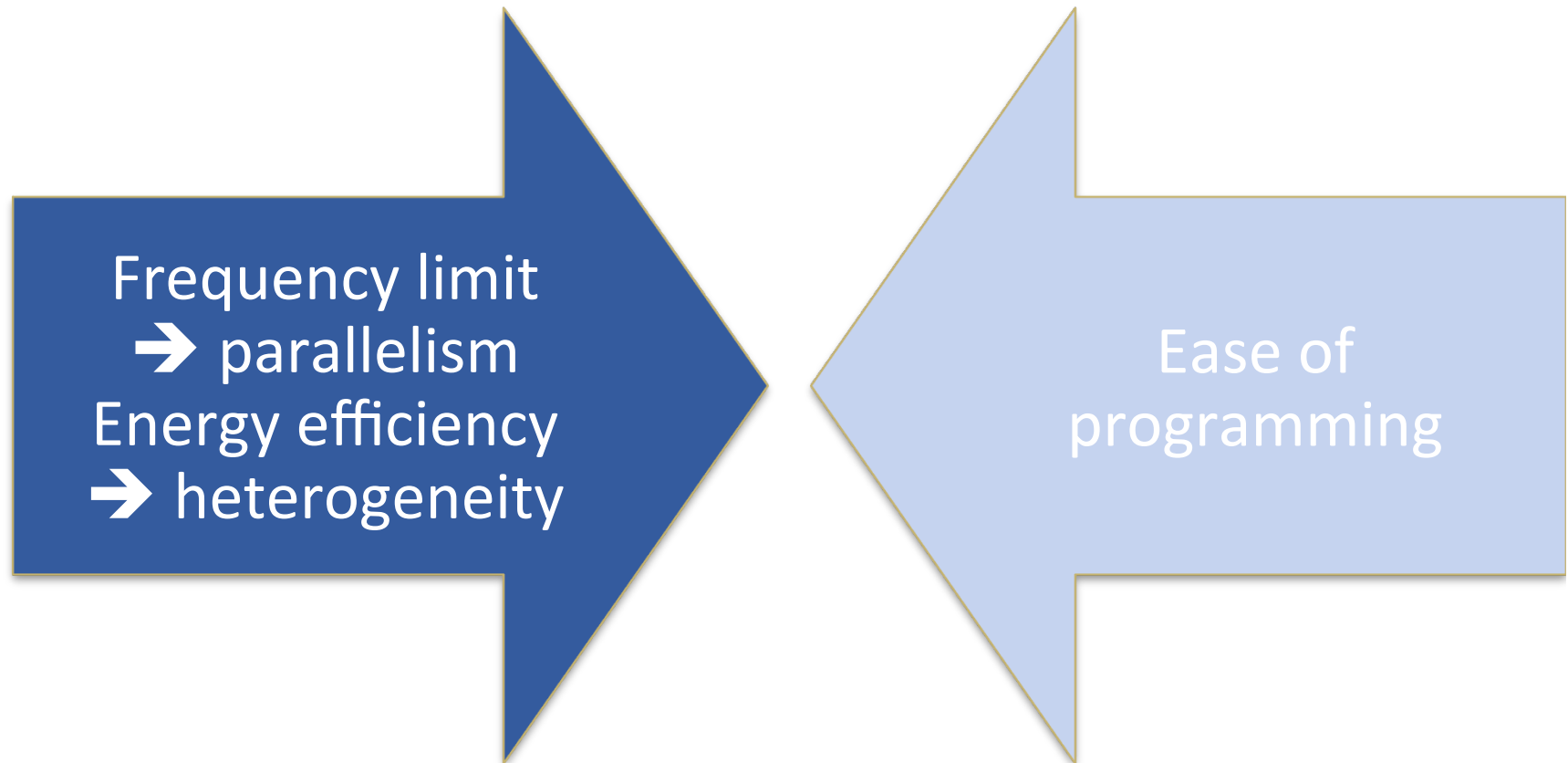
The future of embedded software

ARTEMIS 2006

**Parallelism seems to be too complex for humans ?**

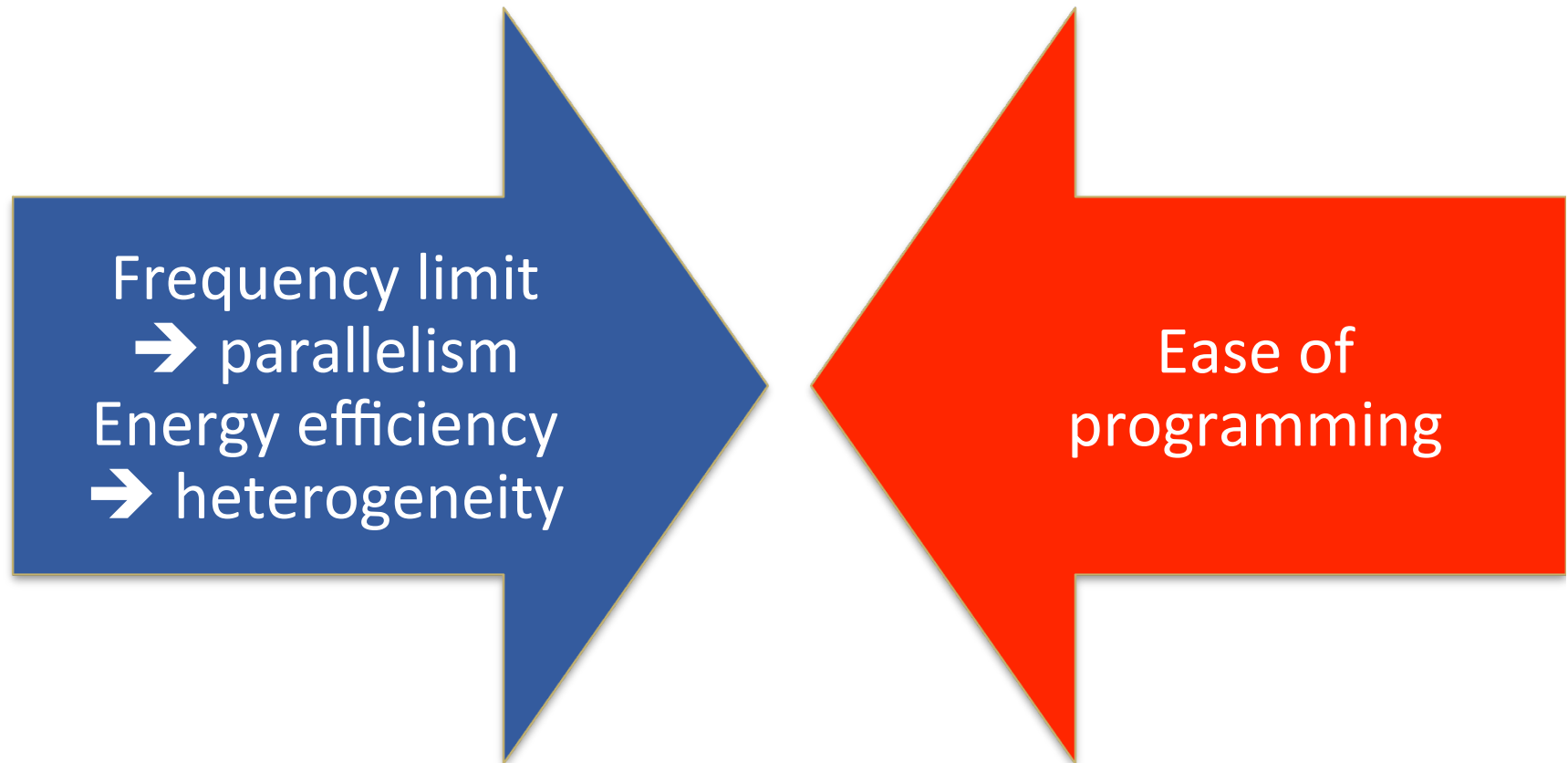


# Parallelism and specialization are not for free...





# Parallelism and specialization are not for free...





# Let the computer do the job

- Describing **what** the program should accomplish, rather than describing **how** to accomplish it as a sequence of the programming language primitives.
- For example, describe the **concurrency** of an application, not how to parallelize the code for it.
- (Good) compilers know better about architecture than humans, they are better at optimizing code



# Hardware design also evolve...

Formal specifications, model-driven design  
Stateflow, StateCharts, LUSTRE, ....



```

#include <stdio.h>
#include <stdlib.h>
#include <sys/types.h>
#include <arpa/inet.h>

void server1(portServ, port)
{
    sockServ2 = socket(AF_INET, SOCK_STREAM, 0);
    struct sockaddr_in monAddr, addrClient, addrServ2;
    socklen_t lenAddrClient;
    if ((sockServ2 = socket(AF_INET, SOCK_STREAM, 0)) == -1) {
        perror("Erreur socket");
        exit(1);
    }
    if ((sockServ2 = socket(AF_INET, SOCK_STREAM, 0)) == -1) {
        perror("Erreur socket");
        exit(1);
    }
    bzero(&monAddr, sizeof(monAddr));
    monAddr.sin_family = AF_INET;
    monAddr.sin_port = htons(port);
    monAddr.sin_addr.s_addr = INADDR_ANY;
    bzero(&addrServ2, sizeof(addrServ2));

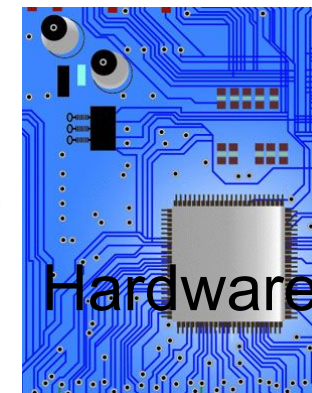
```

Intermediate code  
(C, C++, SystemC...)

Computer processing  
(HLS, core generators  
...), ...)

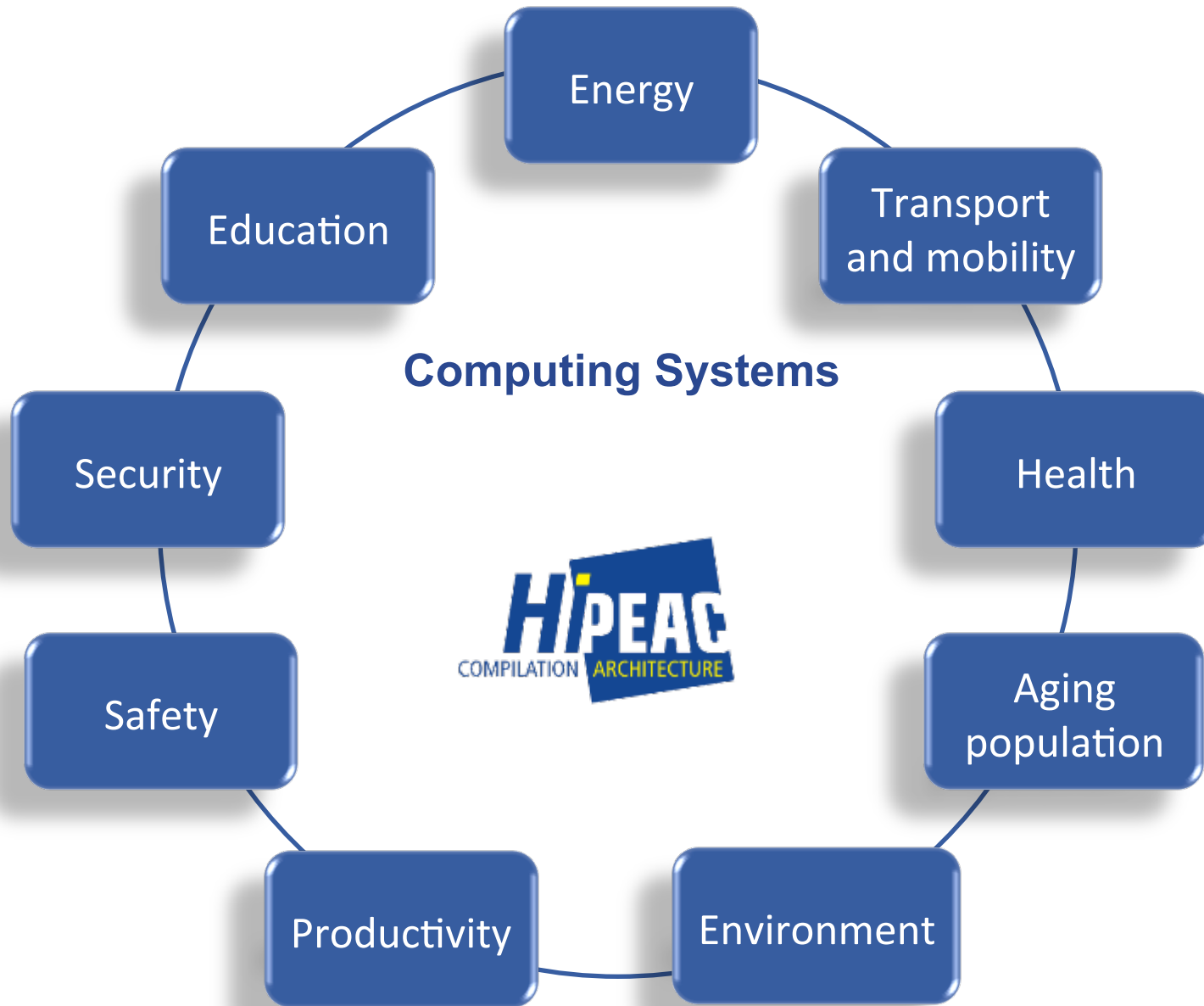


Software



Hardware

# Information technology will bring improved efficiency to solve societal challenges



# Information technology will bring improved efficiency to solve societal challenges

## The Power of 1 Percent

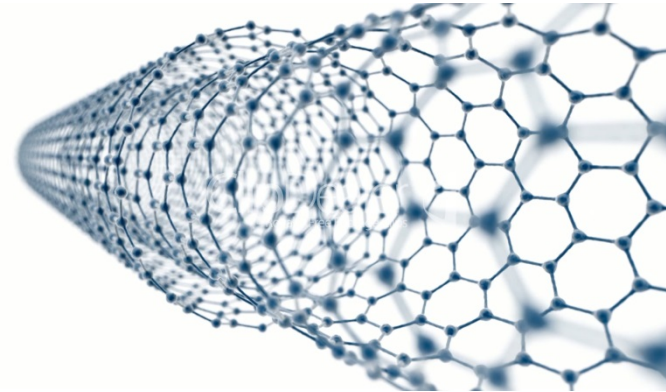
### What if... Potential Performance Gains in Key Sectors

Industry	Segment	Type of Savings	Estimated Value over 15 Years (Billion nominal US dollars)
Aviation	Commercial	1% Fuel Savings	\$30B
Power	Gas-fired Generation	1% Fuel Savings	\$66B
Healthcare	System-wide	1% Reduction in System Inefficiency	\$63B
Rail	Freight	1% Reduction in System Inefficiency	\$27B
Oil & Gas	Exploration & Development	1% Reduction in Capital Expenditures	\$90B

Note: Illustrative examples based on potential one percent savings applied across specific global industry sectors.  
Source: GE estimates



**Conclusion: *What should we do* (as HiPEAC)?**

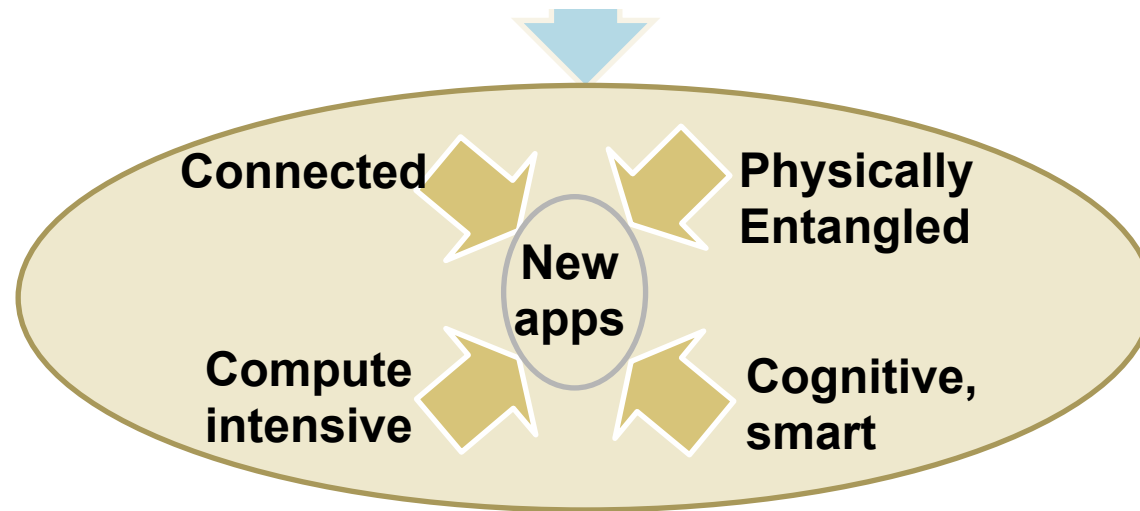


**Multidisciplinary**

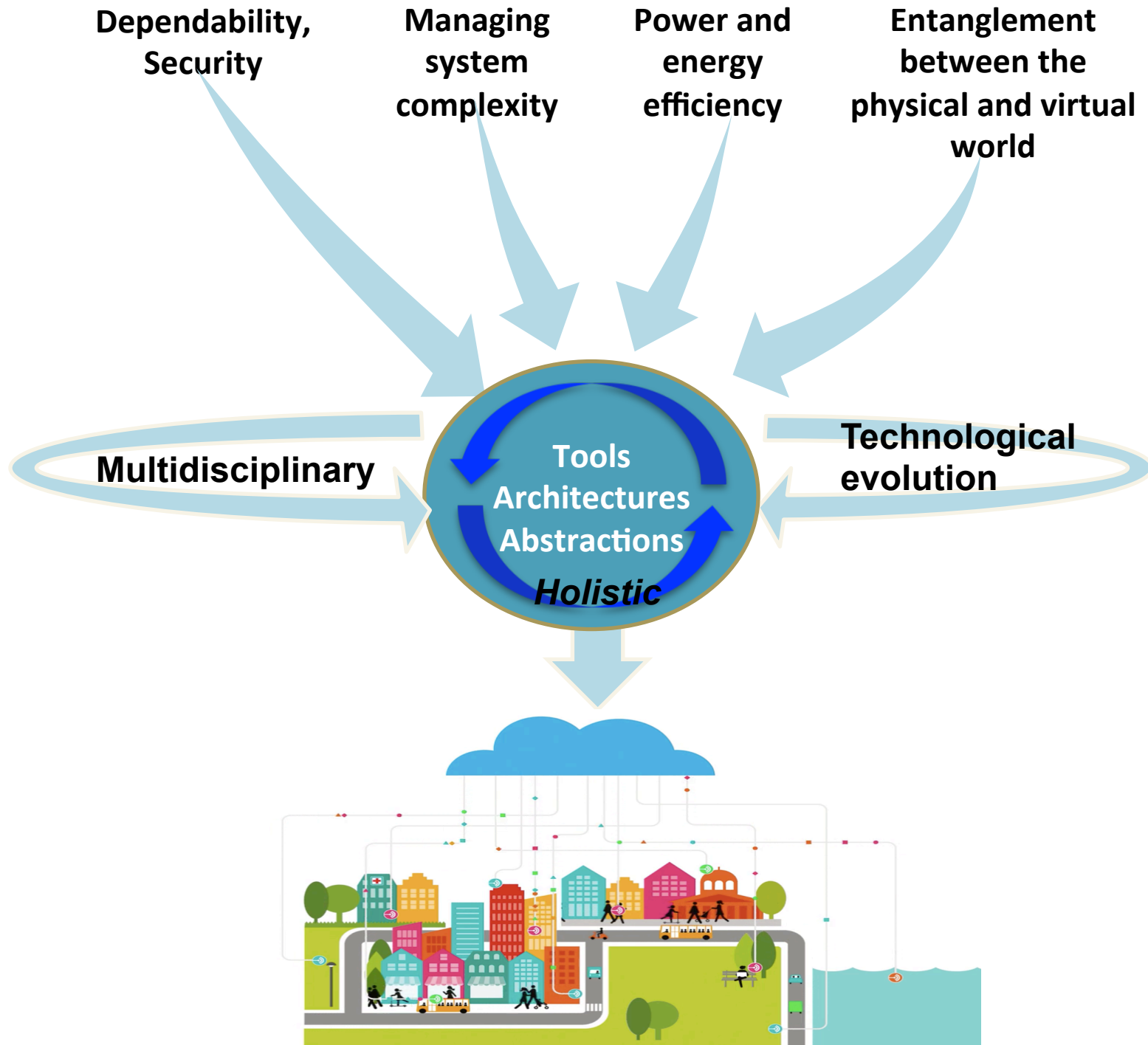
**Technological evolution**

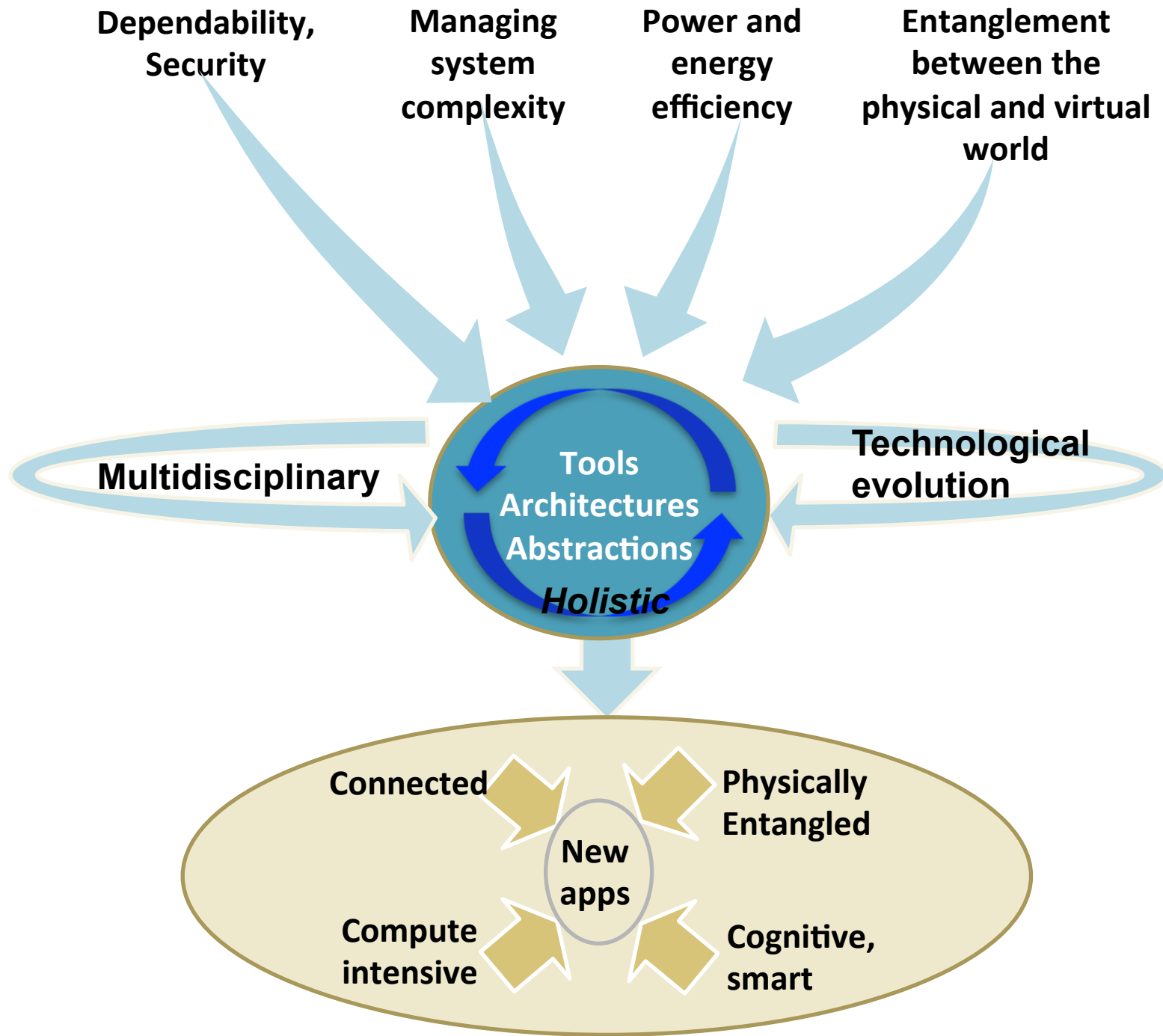
# **HIPEAC**

COMPILATION ARCHITECTURE







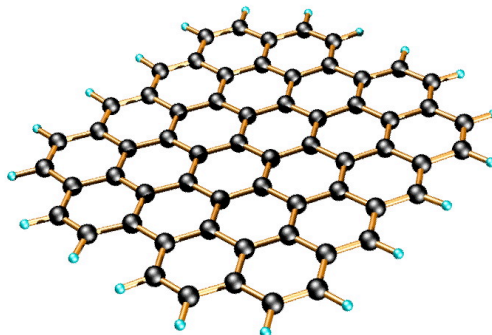
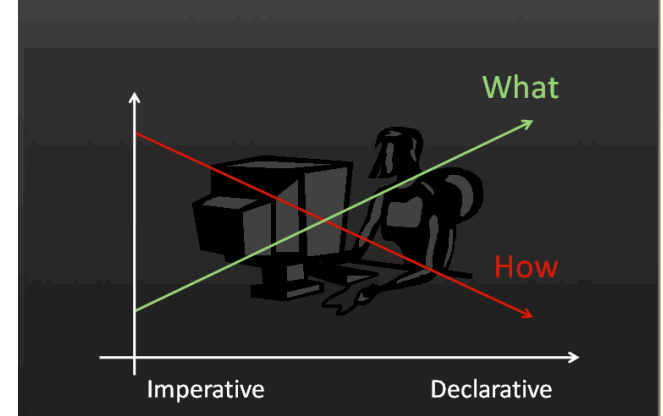


# Time to think differently?

- Approximate computing
- Cognitive computing
- Neuromorphic computing
- Declarative programming
- New computing technologies
  - Graphene
  - Spintronic
  - Quantum...

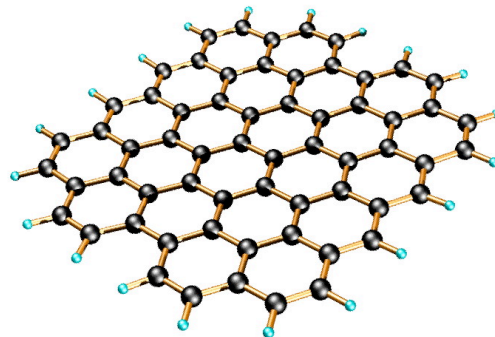
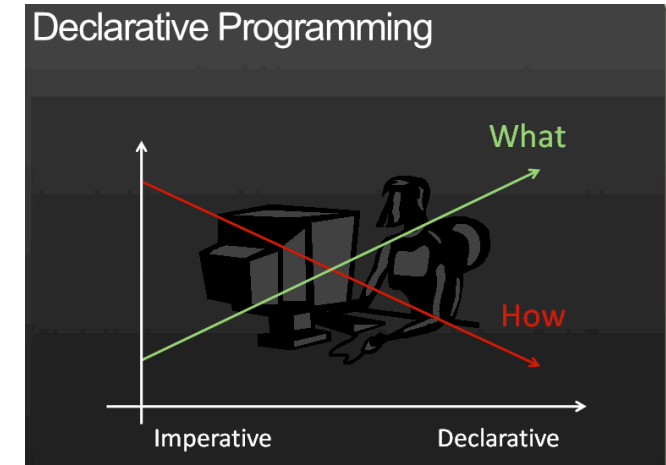


Declarative Programming



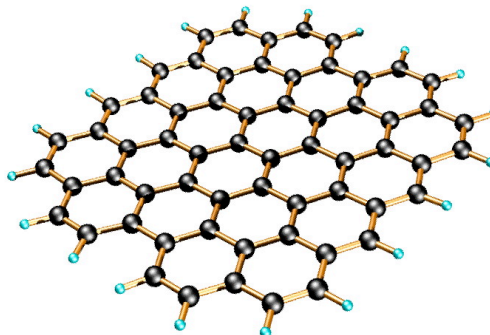
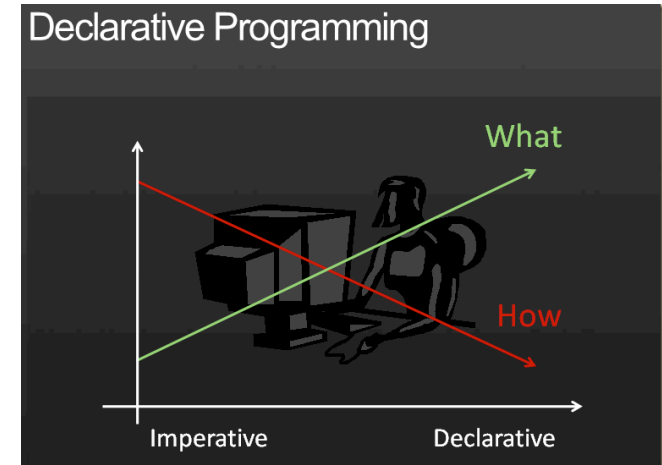
# Time to think differently?

- ~~Approximate~~ computing
- Cognitive computing
- Neuromorphic computing
- Declarative programming
- New computing technologies
  - Graphene
  - Spintronic
  - Quantum...



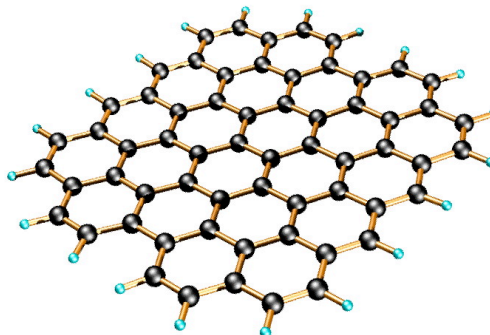
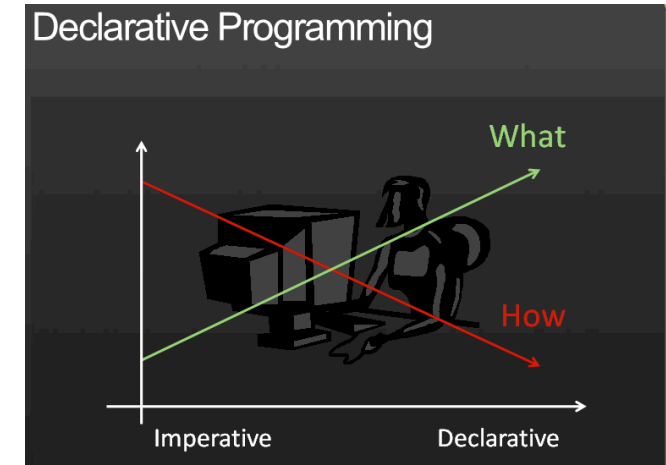
# Time to think differently?

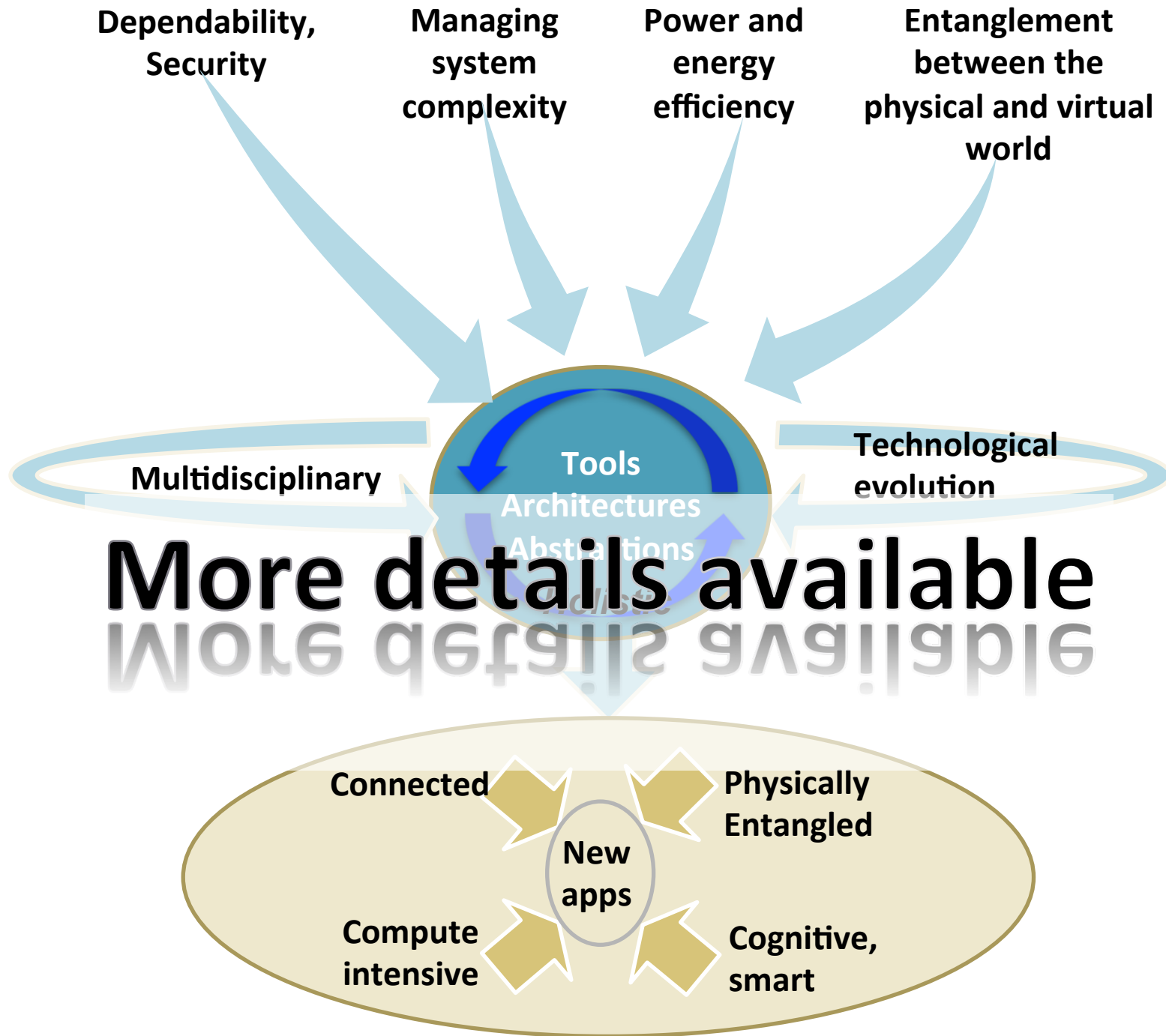
- Adequate computing
- Cognitive computing
- Neuromorphic computing
- Declarative programming
- New computing technologies
  - Graphene
  - Spintronic
  - Quantum...



# Time to think differently?

- Adequate computing
- Cognitive computing
- Neuromorphic computing
- Declarative programming
- New computing technologies
  - Graphene
  - Spintronic
  - Quantum...







Dependability  
Security

Integration  
of the physical  
and digital world

Multidisciplinary

Technological evolution



# HiPEAC Vision 2015

HIGH PERFORMANCE AND EMBEDDED ARCHITECTURE AND COMPILATION

Editorial board: Marc Duranton,  
Koen De Bosschere, Albert Cohen,  
Jonas Maebe, Harm Munk

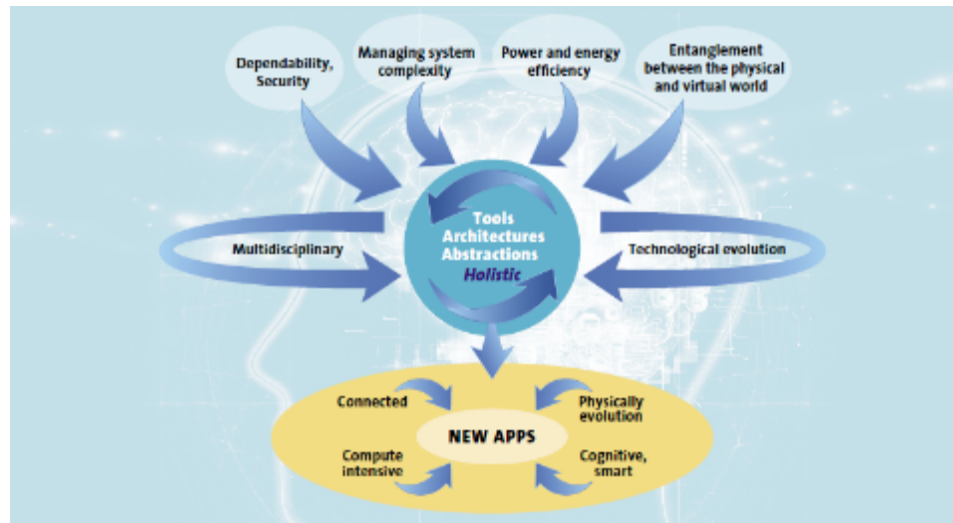
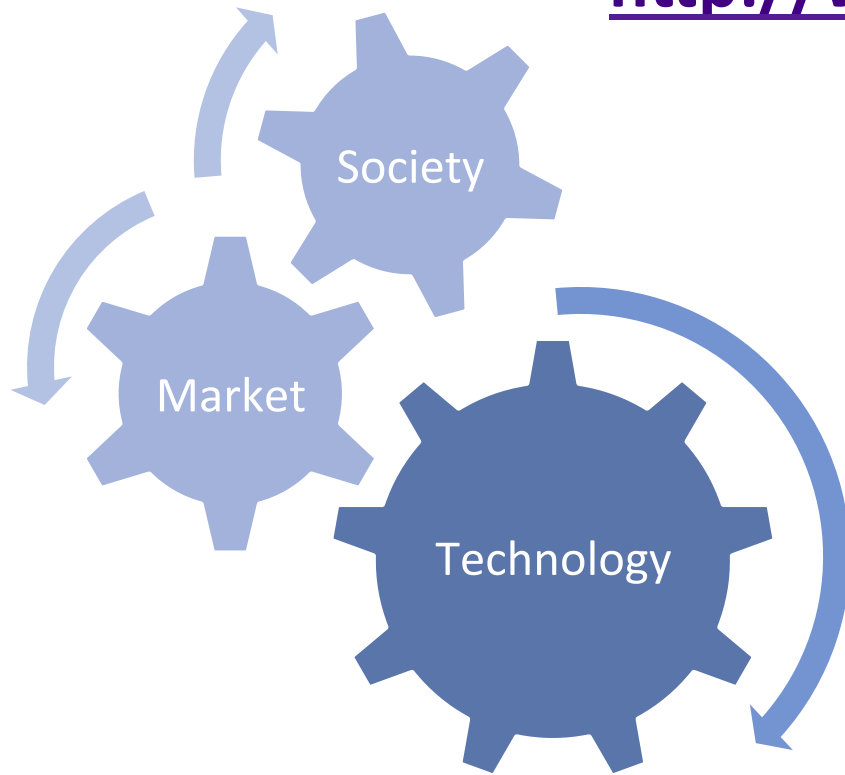


<http://www.hipeac.org/vision>





<http://www.hipeac.org/vision>



© 2015 HIPEAC