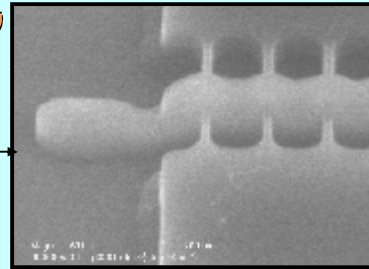


Ambient Intelligence: Giga-Scale Dreams and Nano-Scale Realities



Giga-Scale Complexity



Nano-Scale Realities

Hugo De Man
Prof. Emeritus K.U.Leuven
Senior Fellow IMEC
deman@imec.be

The Ambient Intelligence Dream



**Secure, trustworthy computing and communication
embedded in every-thing and every-one.**

**A pervasive, context aware ambient,
sensitive and responsive to the presence of people**



Aml = Interplay of 4 Technologies

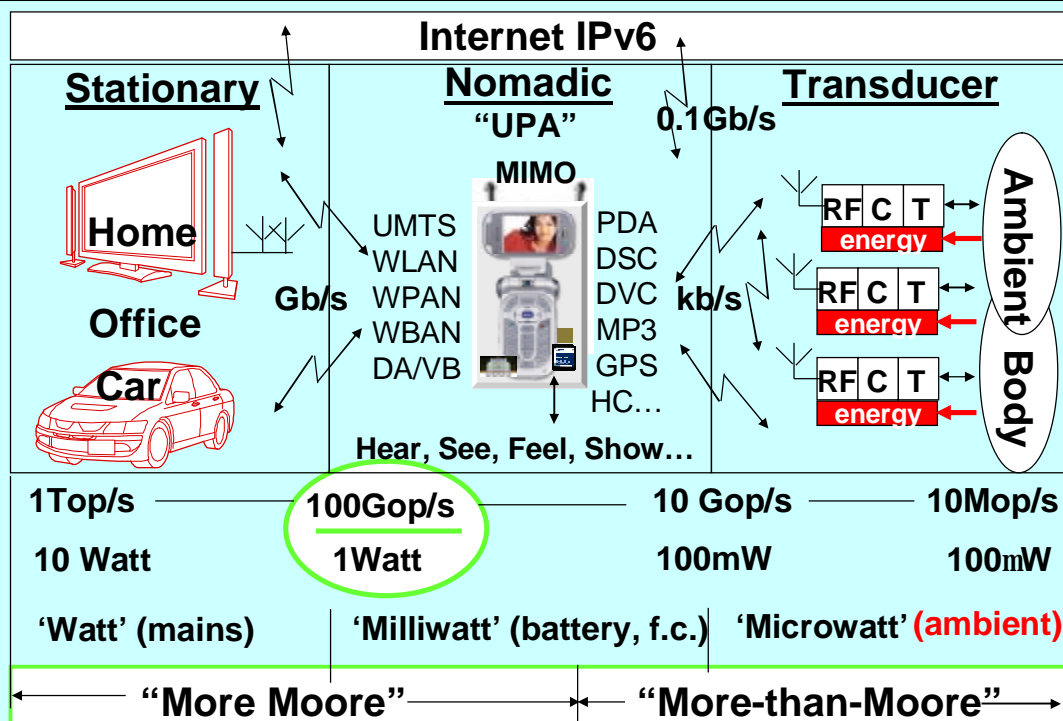


- Ø Embedded computing
- Ø Ubiquitous wireless communication
- Ø Transducer networks i/f to physical world
- Ø Multi-mode, natural interfaces to the user

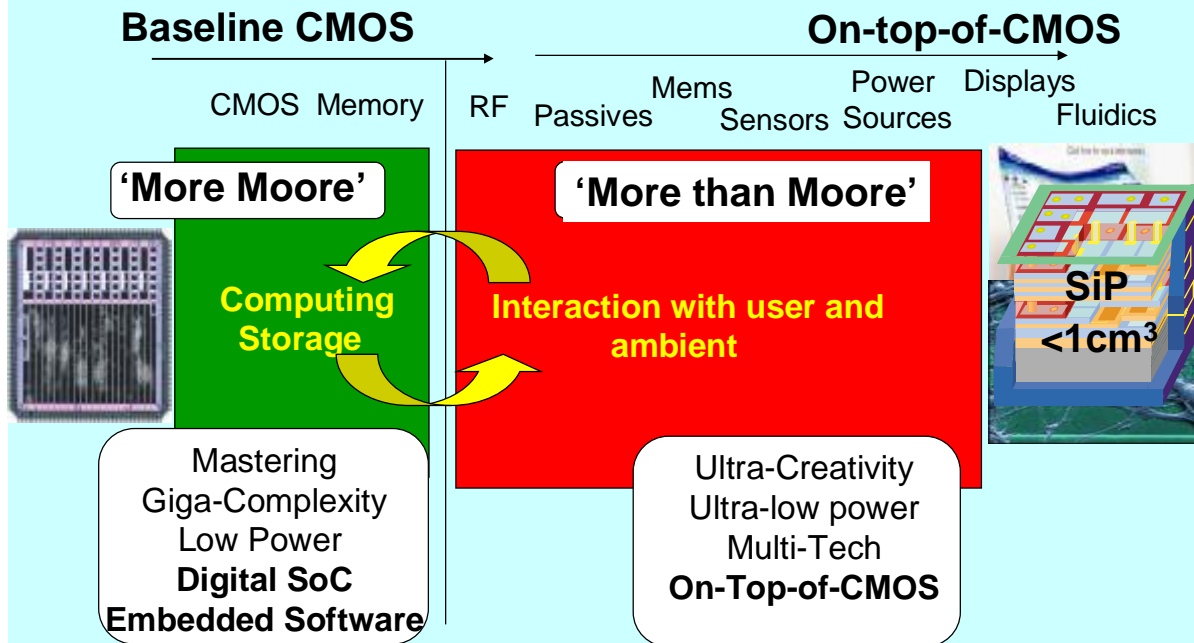


©ISTAG

Device Classes for Aml



'More Moore' and 'More than Moore'



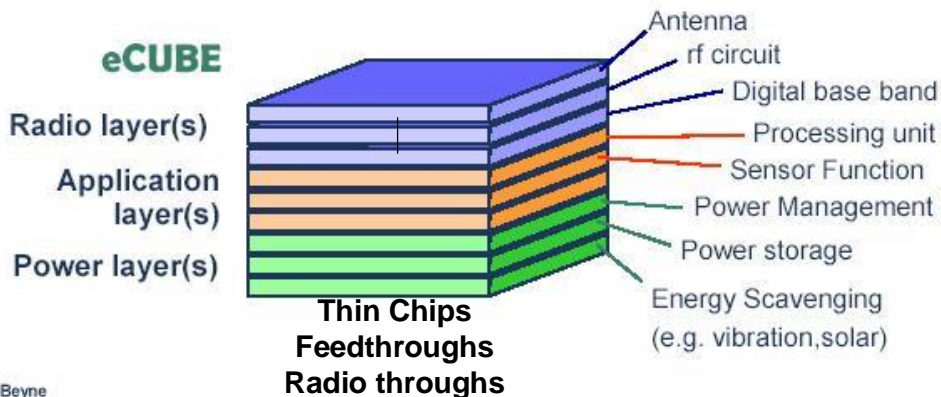
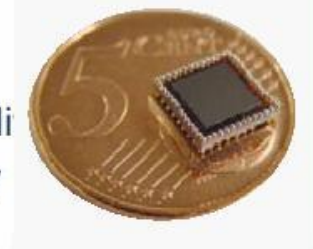
Source: Medea+



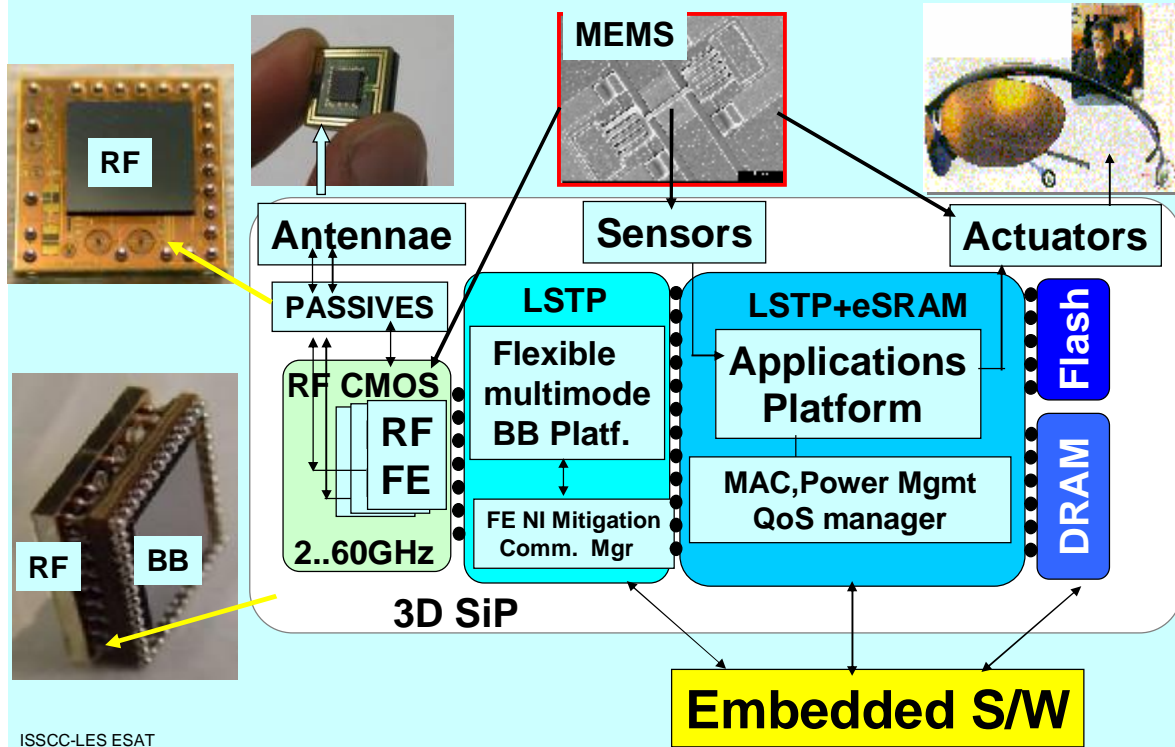
3D Integration technologies

3D-SIP

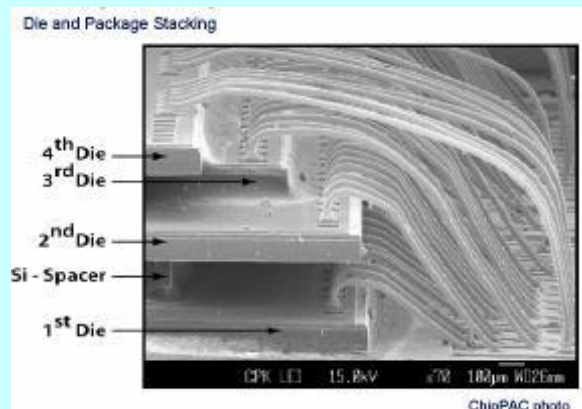
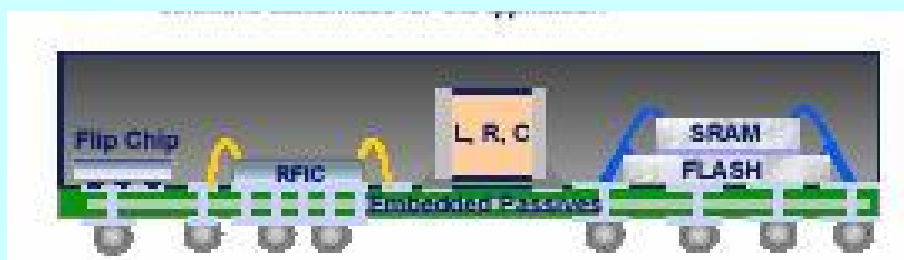
- Stacking of SIP sub-systems
- Excellent yield and manufacturability
- Relatively low 3D interconnectivity
- Realisation of "e-Cubes"



UPA = 'More Moore' + 'More than Moore'



SIP integration



Philips
TI
Nokia

...

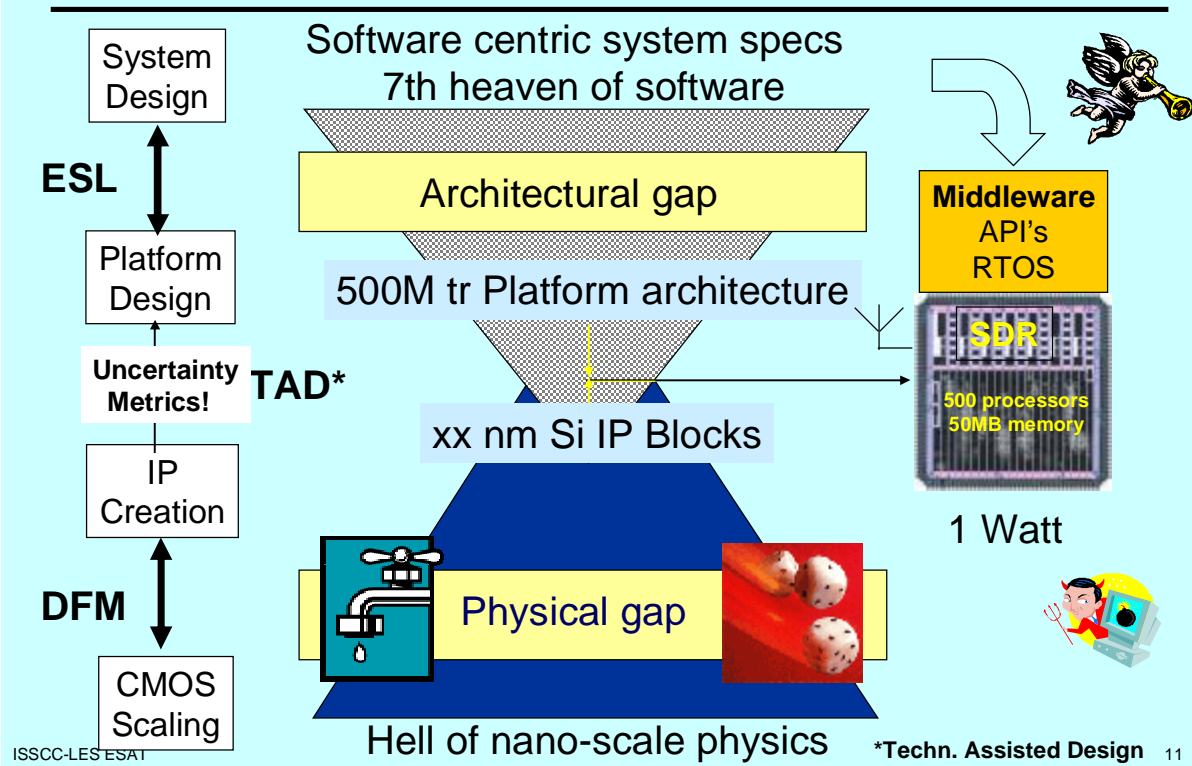
Outline

- **The Ambient Intelligence Dream**
- **More Moore: Managing Giga-Complexity**
Ø of digital Watt and MilliWatt SoC's
- **More than Moore: Ultra-Creativity**
Ø for ultra-low-power and -cost interfaces
- **Conclusions**

Challenges for Digital Aml

- **Power Efficiency (PE): 10 to 200 GOPS/Watt**
- **Part cost < 10 \$ NRE cost > 50M\$**
Ø PE two-orders-of-magnitude > GP microprocessor
at 1/20 th of part cost
- **Adaptive by embedded software**
Ø From Hw centric ASIC to Sw centric PLATFORM
- **Real-time stream based processing (MM, SDR)**
Ø Memory intensive (> 70% area, power)

Gaps between Aml-Dreams and Nano-Scale Realities...



1. The Architectural Gap

Create methods, tools and skills
for designing
flexible, yet power-efficient platforms
and API for mapping real-time
MULTIPLE S/W applications on them

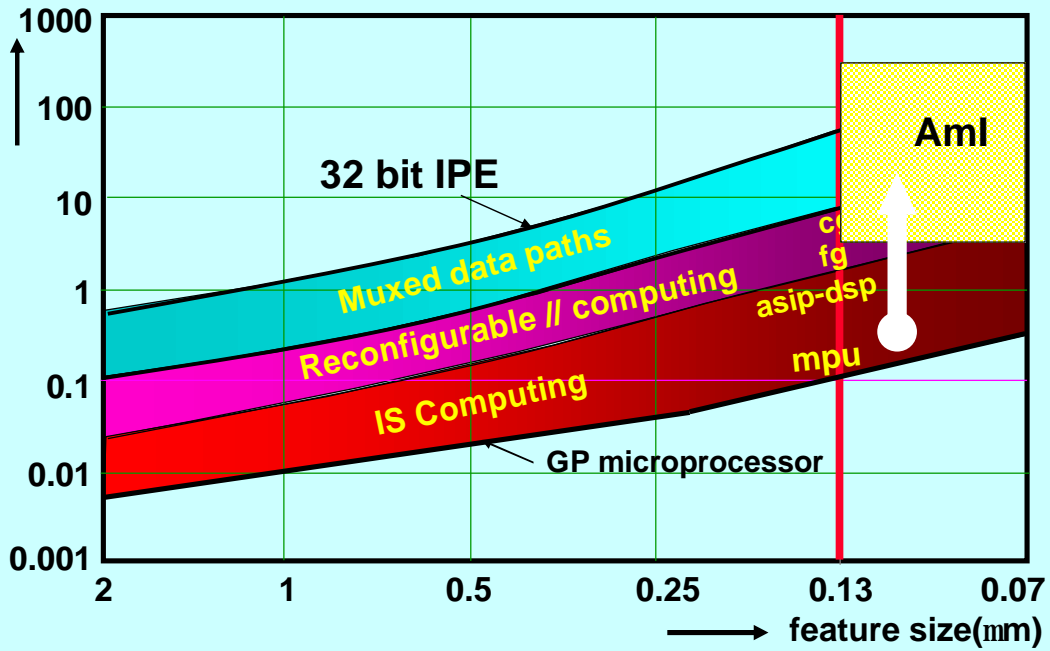
Ø Flexible: programmability and/ or reconfigurability

Greatest Challenge:
Overcome Power-Flexibility Conflict

Power-Flexibility Conflict

Power efficiency PE (GOPS/Watt)

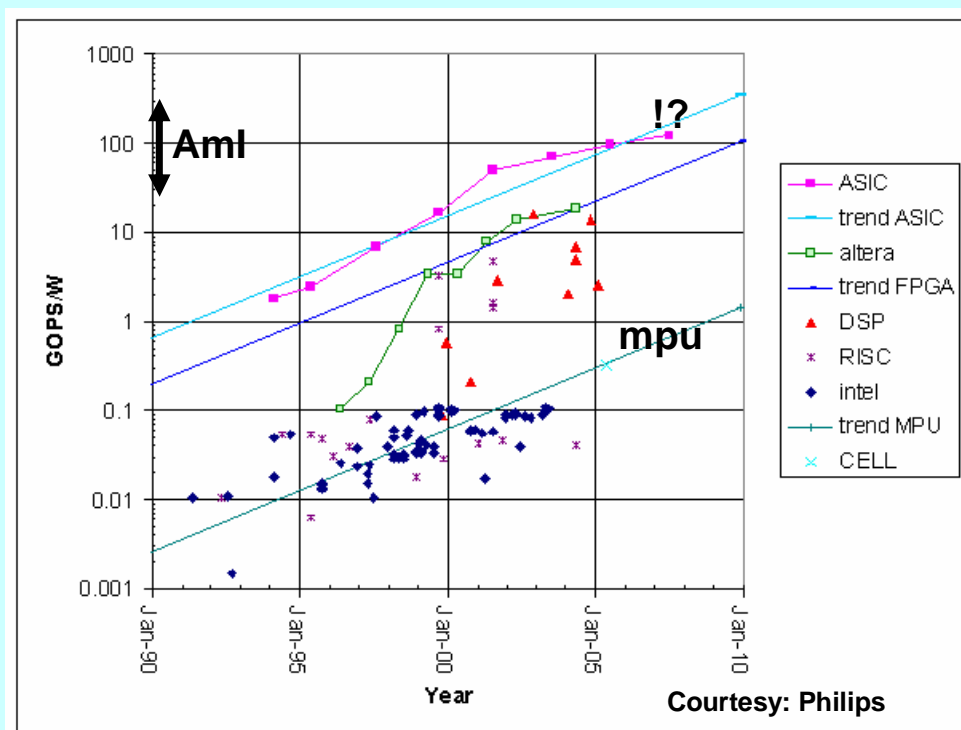
Source: T.Claesen (ISSCC99)



ISSCC-LES ESAT

13

Power-Flexibility conflict in practice



ISSCC-LES ESAT

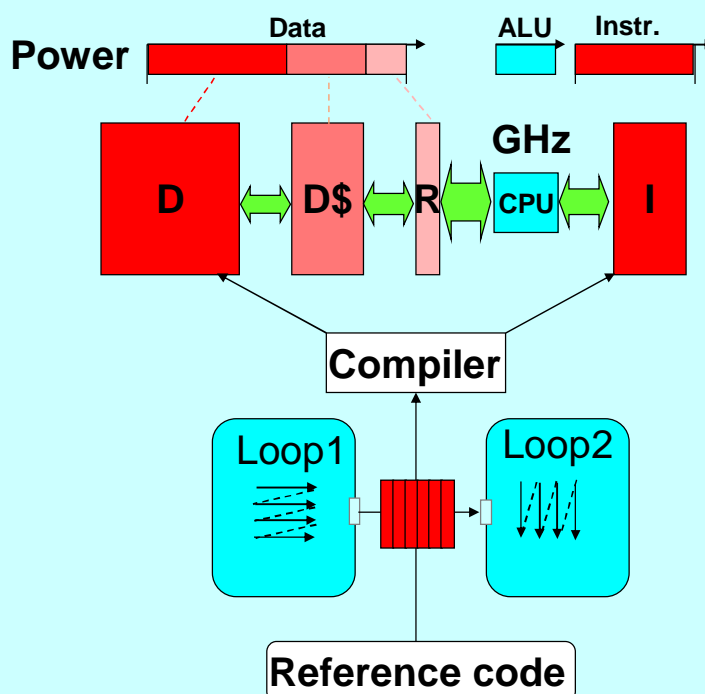
Courtesy: Philips

14

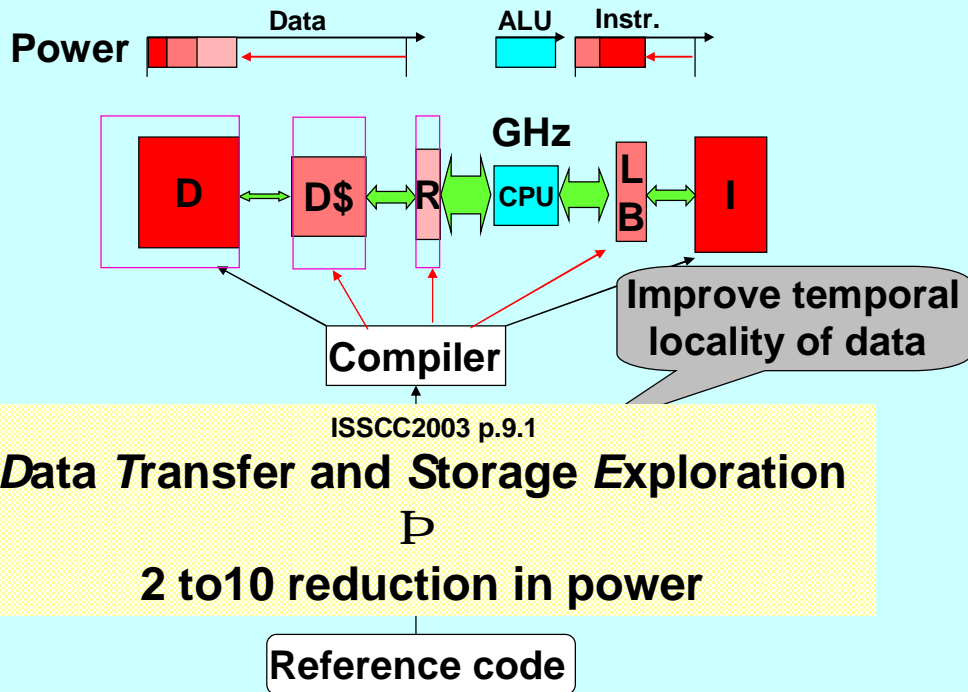
How to Reconcile PE and Flexibility?

- **Cleaning software for low power**
 - ∅ Minimize # memory transfers, operations = energy
 - ∅ Regularize, vectorize code, min. data-size
 - ∅ Factor 5 to 10 lower power but **tools and skills** needed
- **Exploit parallelism to reduce power**
 - ∅ Better to use 20 processors @ 100MHz than 1 @ 2Ghz
 - ∅ But beware of nano-scaling (see later)
 - ∅ Urgent need for design **methods, models, tools and skills** to support Hw/**Sw** co-design (devil is in the software!)
- **Use low power task specific processors**
- **Exploit task level dynamism for sub 1 Volt op**
 - ∅ Just enough V_{DD} , F_{cl} , I_L

Overcome the Memory Wall...

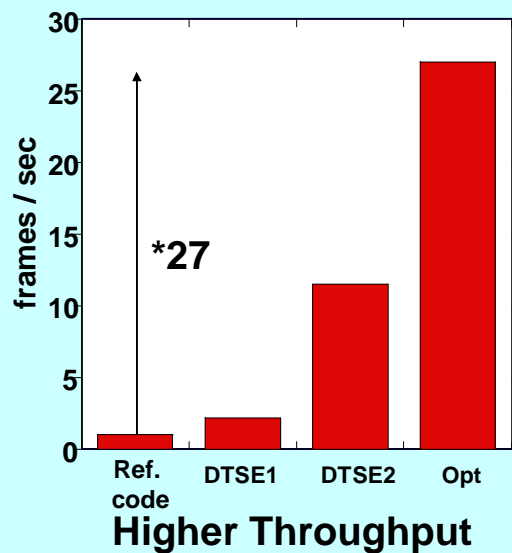
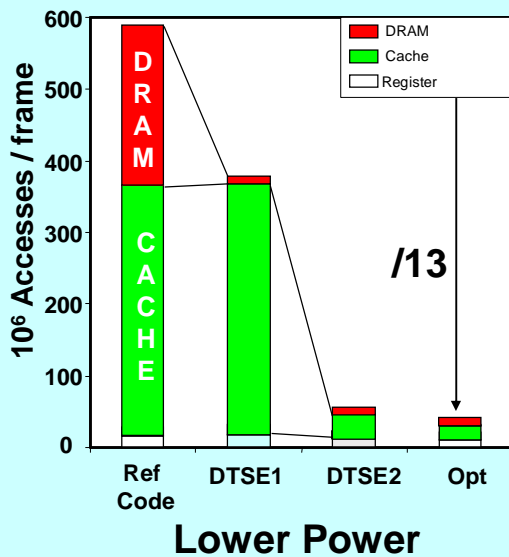


Cleaning Software for Low Power



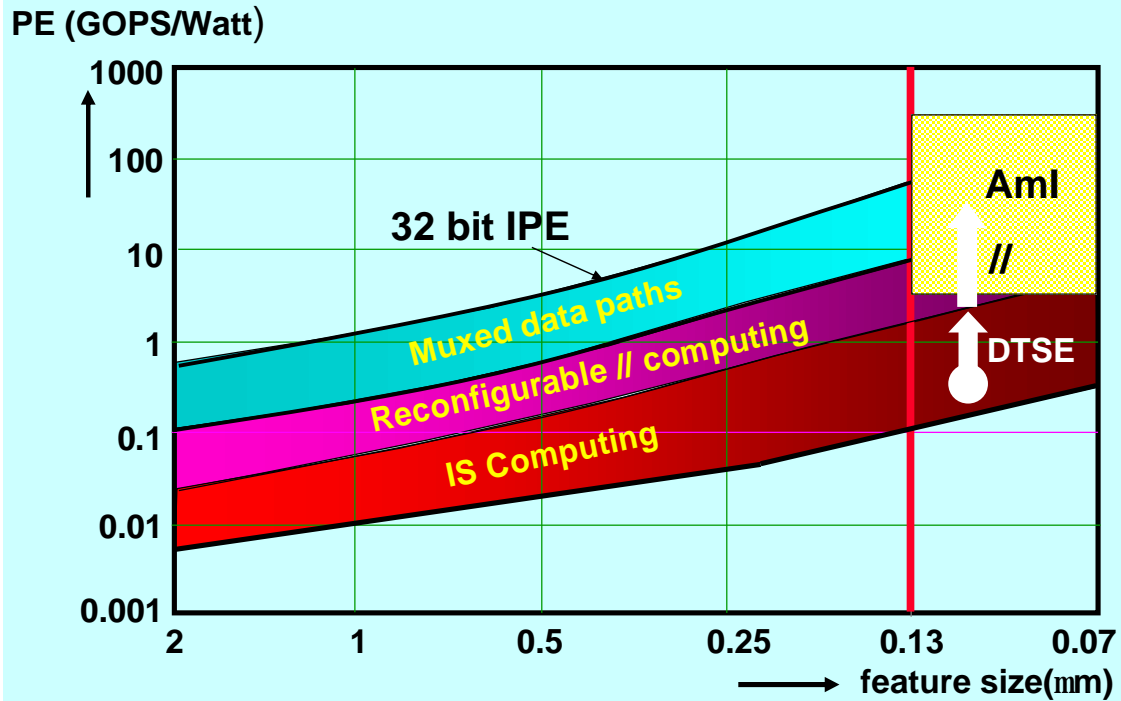
Example: Cleaning Software for Low Power

VGA quality MPEG 4 encoding on 1.6 GHz Pentium M



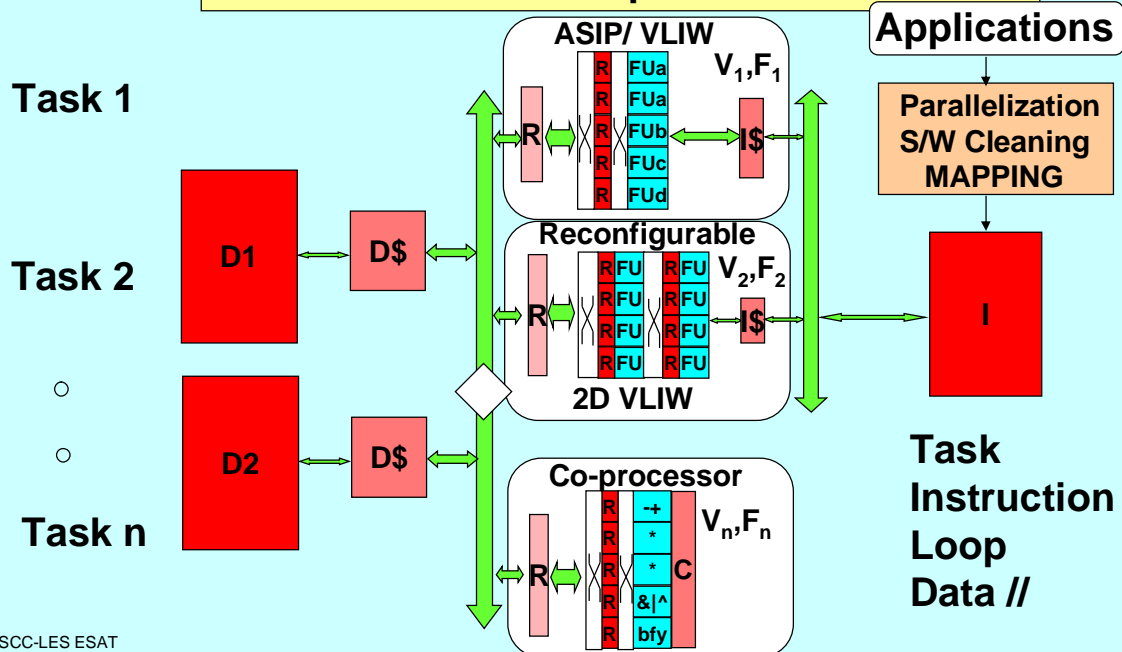
Tools available (www.powerscape.com) but far from solved for parallel / multi-tasking architectures !

DTSE Improves PE by 2 to 10



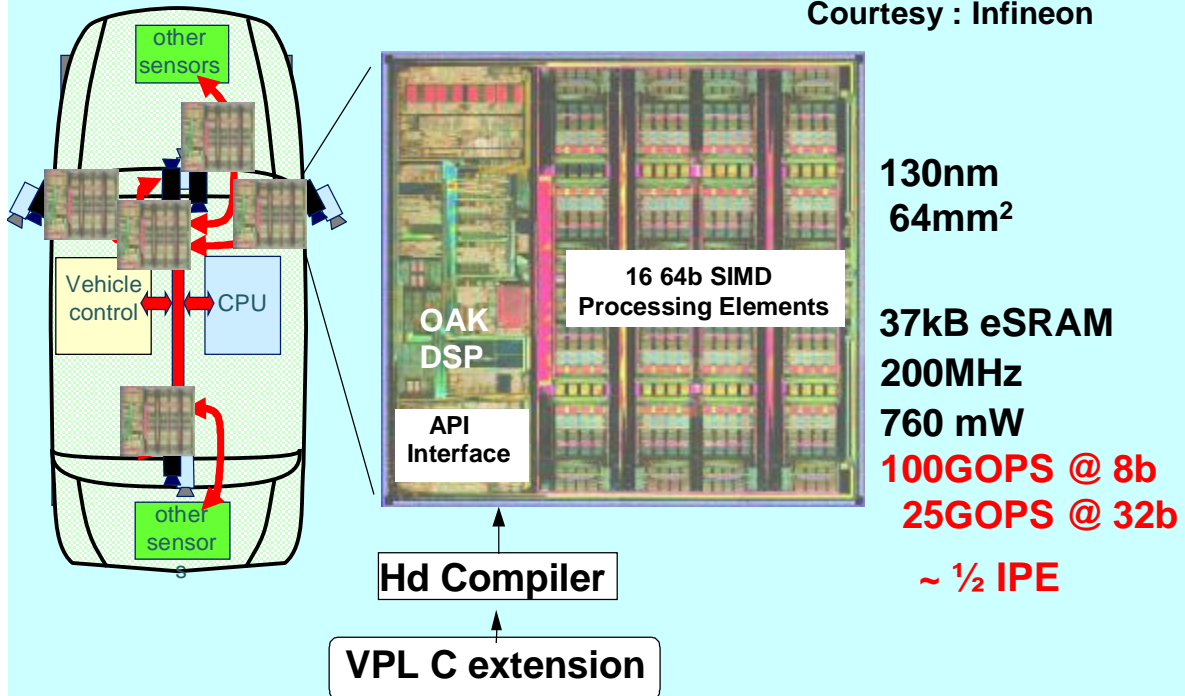
Use Low Clock Speed Parallel Architectures

Spatial locality of storage and computation leads to low power



Vision Platform (VIP)

Courtesy : Infineon

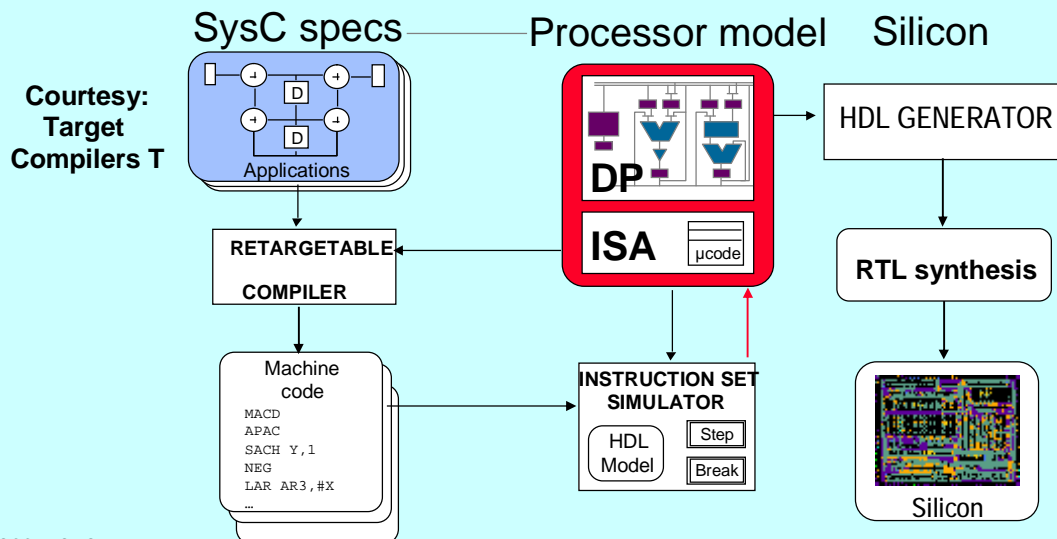


ISSCC-LES ESAT

21

Task Specific Processor Design...

RWTH AACHEN ▷ Lisatek(CoWare);
IMEC ▷ Target Compiler T, ARM Optimode
PHILIPS ▷ Siliconhive; TENSILICA, PicoChip...

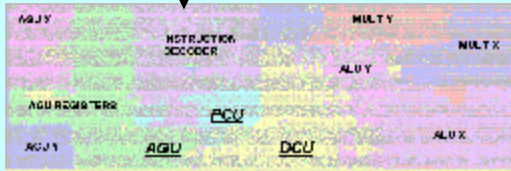


ISSCC-LES ESAT

22

Co-Design of Processor AND Compiler

Retargetable C compiler

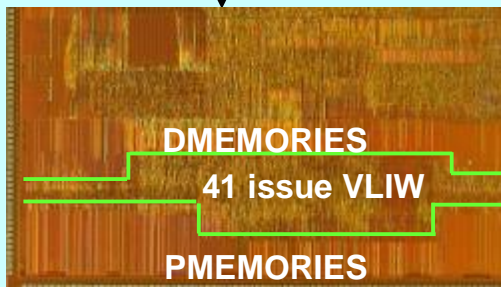


Courtesy: Philips-Target Compilers

Coolflux Audio ASIP

130 nm 0.9V 0.32mm² 24bit
 2.0 mW MP3 incl. SRAMs
42 MOPS/mW (~1/4 IPE)

Retargetable C compiler



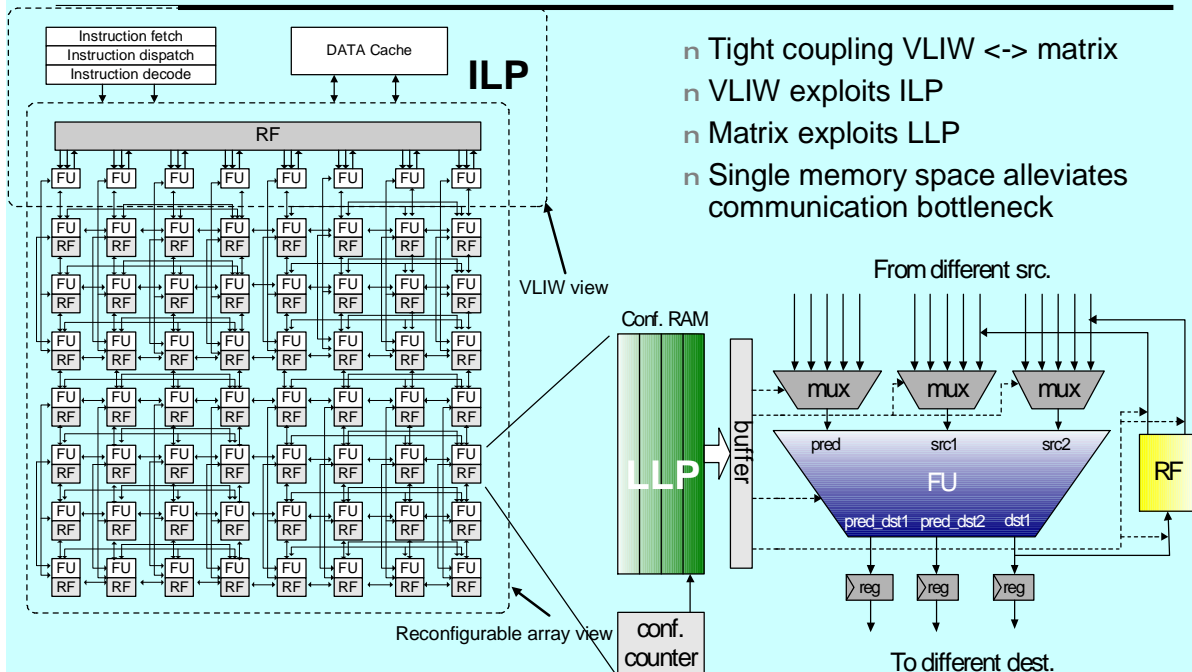
ISSCC-LES ESAT Courtesy: SiliconHive

41 Issue SDR VLIW

130 nm 1.2V 6.5mm² 16 bit
 30 operations / cycle (OFDM)
 150 MHz 190mW (incl SRAMs)
24 GOPS/W (~ 1/5 IPE)

23

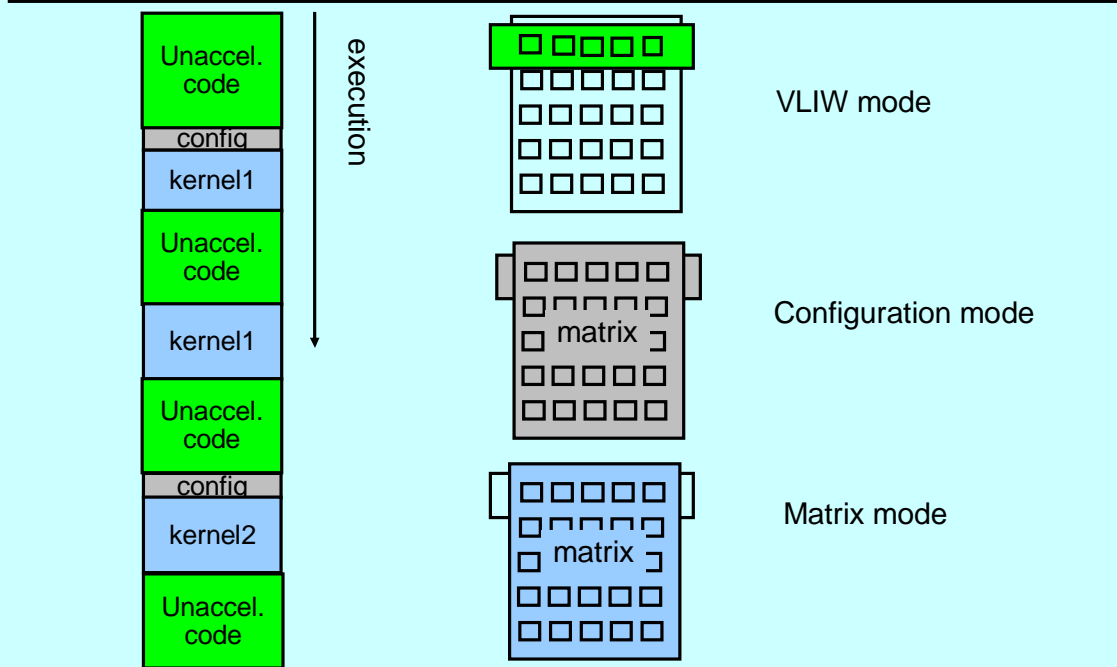
2-D Low power coarse grain core (IMEC ADRES Template Architecture)



- n Tight coupling VLIW <-> matrix
- n VLIW exploits ILP
- n Matrix exploits LLP
- n Single memory space alleviates communication bottleneck

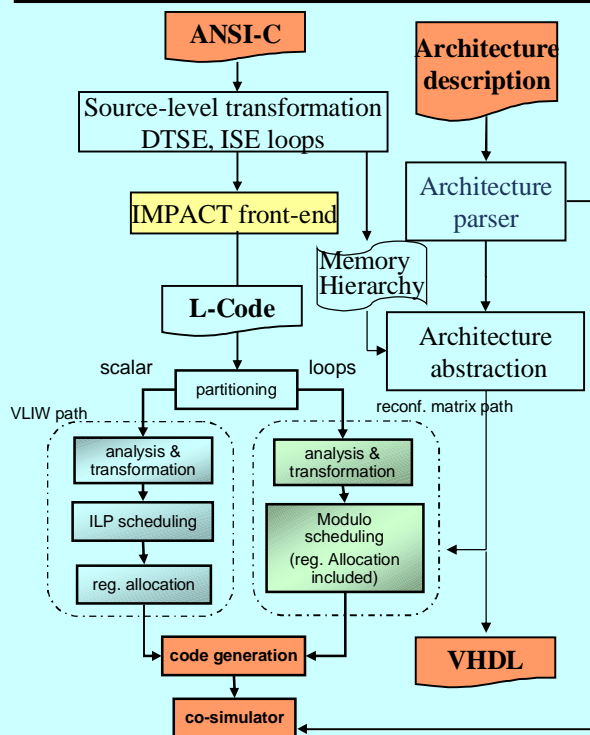
24

Execution Model of ADRES



Execution of VLIW and the matrix are alternated

ADRES is supported by retargetable mapping tools



DRESC (*Dynamically Reconfigurable Embedded Systems Compiler*):

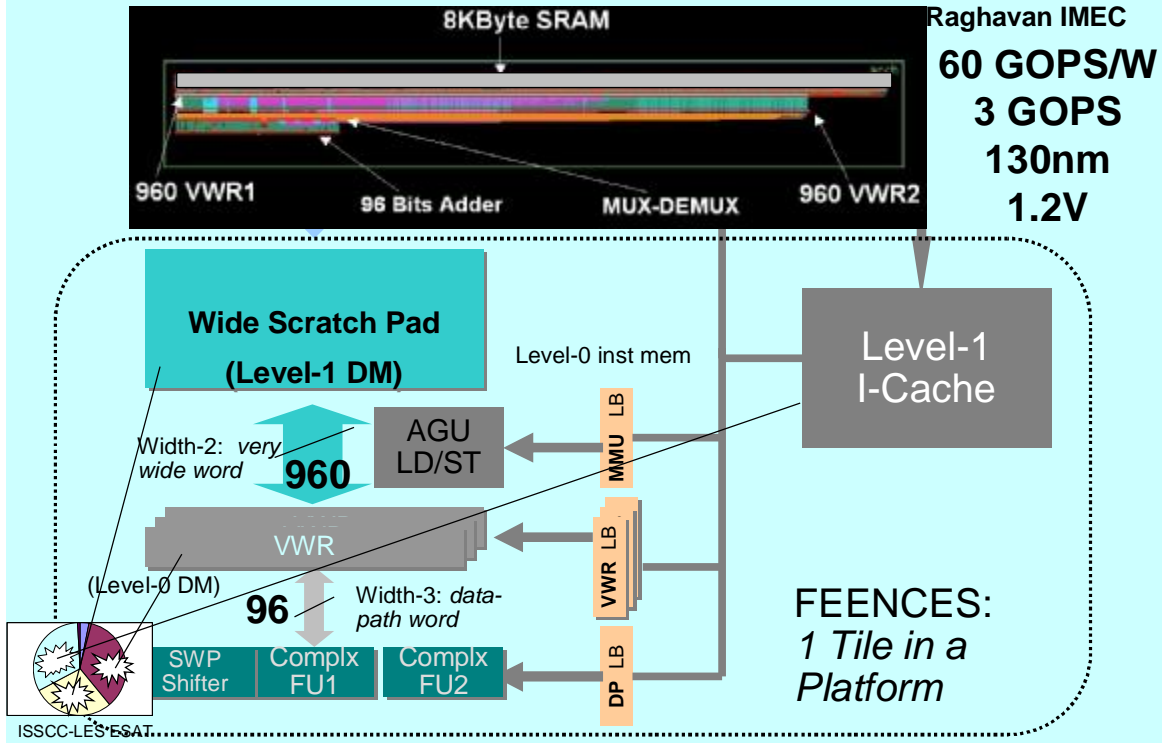
Bingfeng Mei et al, ICFPT2003

Modulo scheduling:

Bingfeng Mei et al, DATE2003

**Architecture + Compiler
Fully Operational on FPGA
for 4*3 (MPEG 2, IDCT...)
8*8: 300MHz, 4mm², 250mW, 90nm
1.2 Volt
10 GOPS
40 GOPS/Watt (est)**

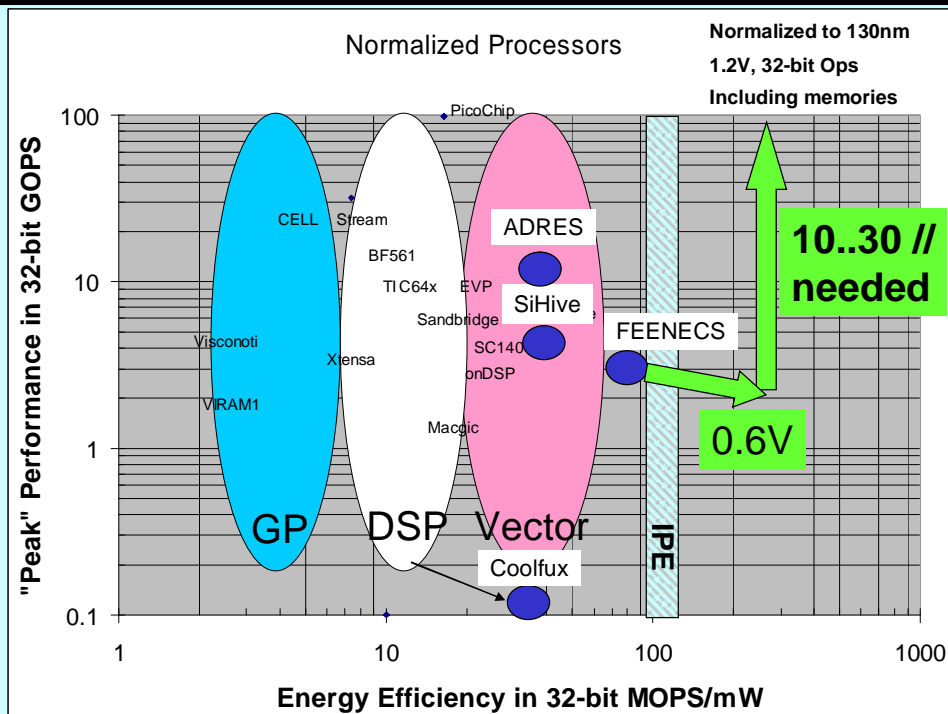
Other IMEC Experiment FEENECS Very Wide L1, REG



27

“Estimated” State of the Art Processors

Courtesy: Raghavan IMEC

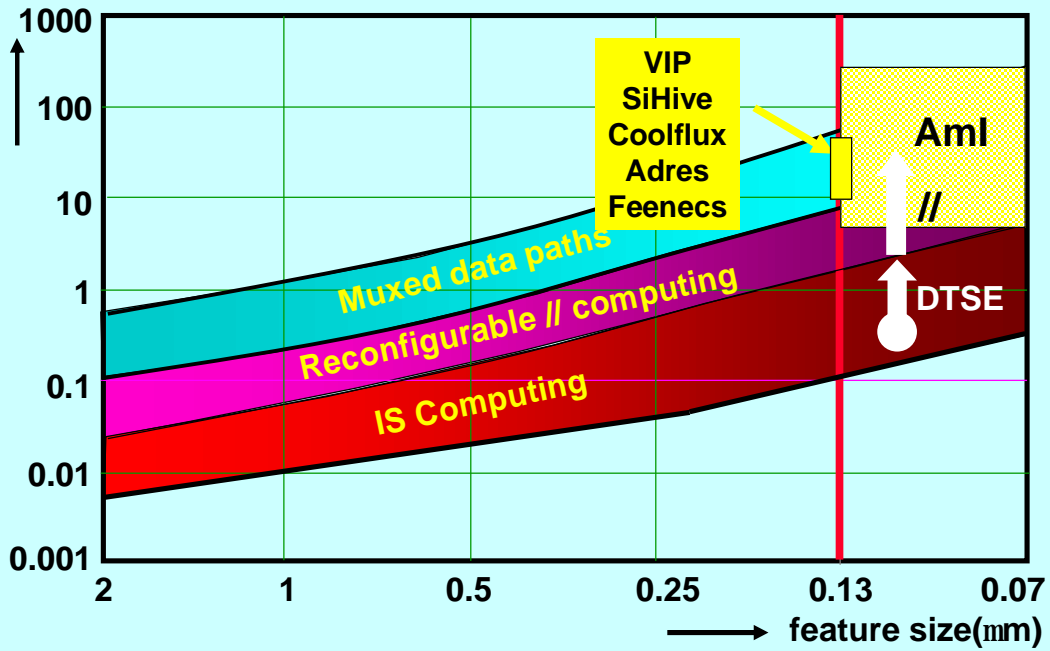


ISCC-LES ESAT

28

All Close to IPE and Flexible...

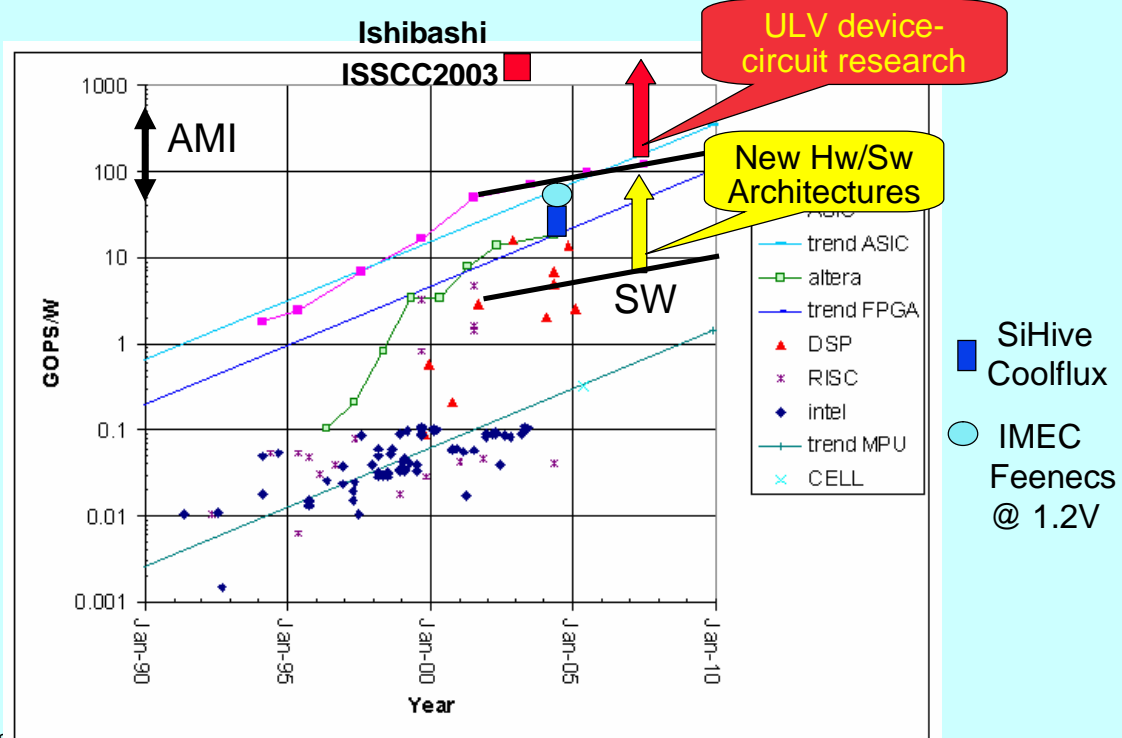
PE (GOPS/Watt)



ISSCC-LES ESAT

29

If we want more...

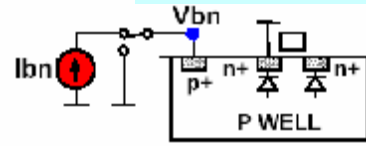
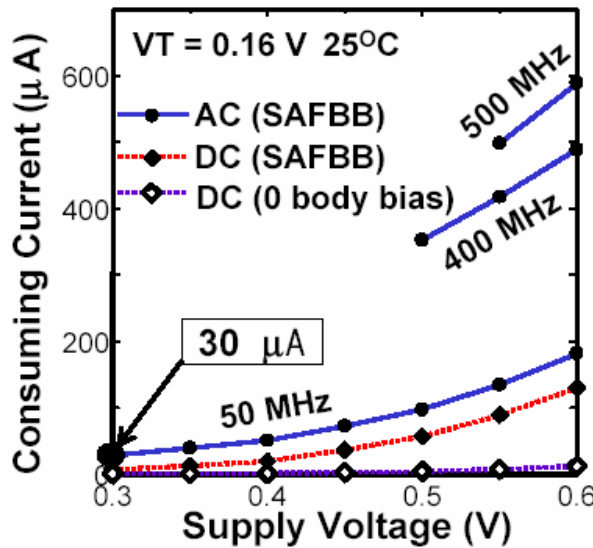


ISS

30

Example: Ishibashi ISSCC2003

Consuming Current of 32-bit Adder

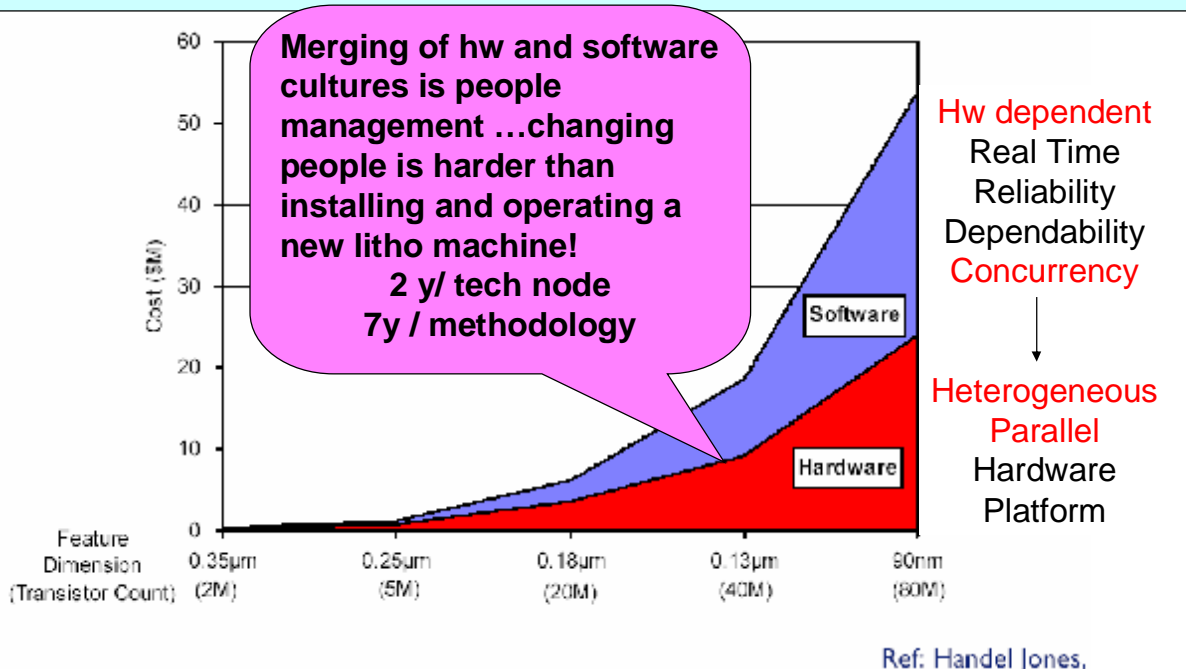


130nm
SAFBB
1400Mops/mW @ 500MHz
5 Mops/µW @ 50 MHz

Parallel Architectures
Novel Compilers @ OS
Managing Variability

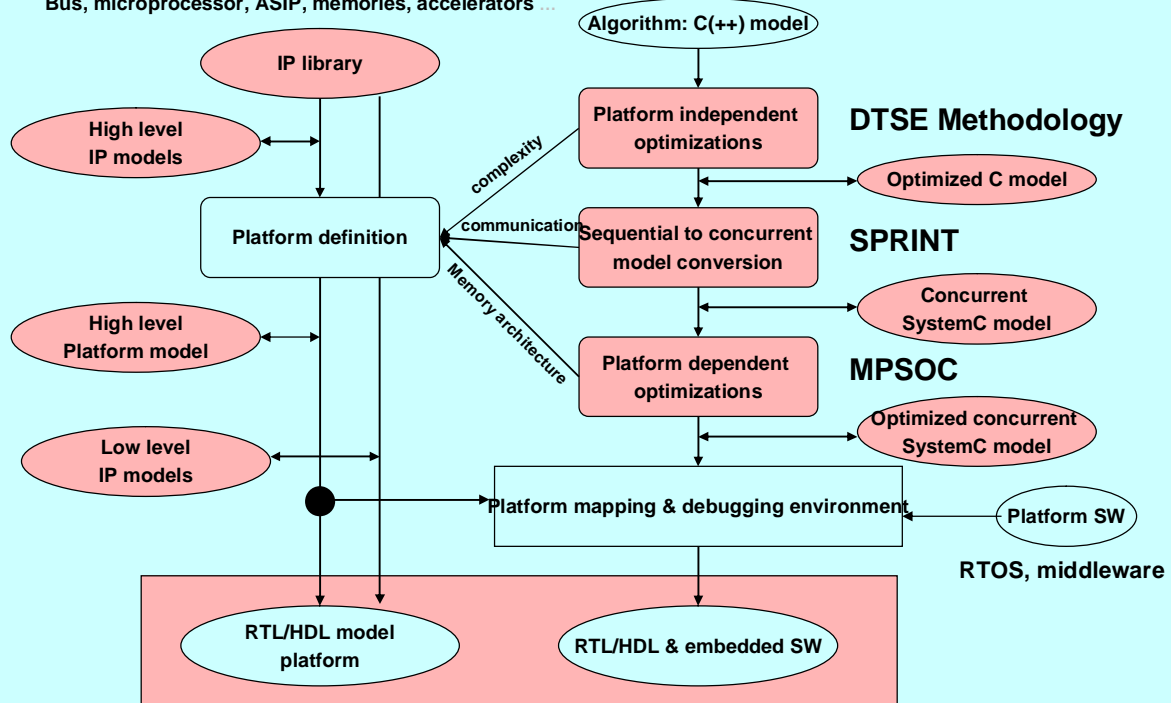
Achilles' Heel #1: Software Cost

Software productivity increase 5.8%/y <-> 58%/y Moore



In urgent need: multi-core design flow

Bus, microprocessor, ASIP, memories, accelerators ...



Achilles' Heel #2: Communication...

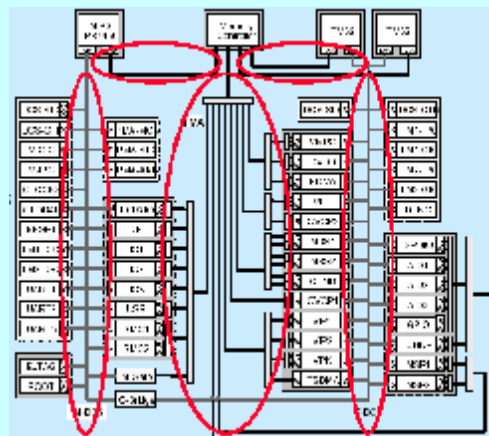
NEXPERIA DVP platform

UHAPI



130nm 50M tr, 250MHz
1 MIPS, 2 Trimedia
60 IP blocks, **250 RAM's**
Up to 100 Gops, **4 Watt**

Bus Based Platform



Buses don't scale
And physics does not help either

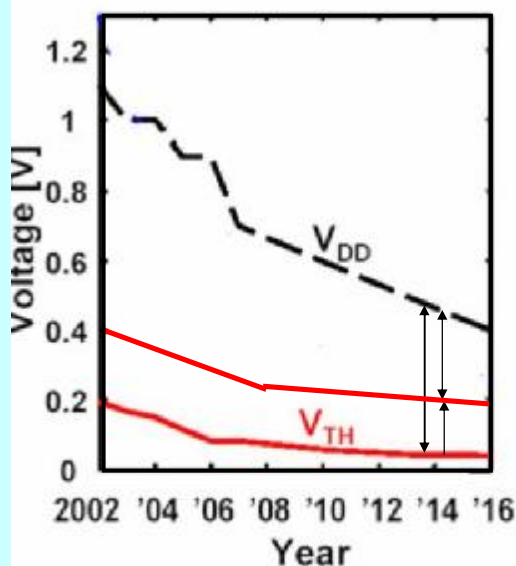
2. Hell of Nano-Scale Physics



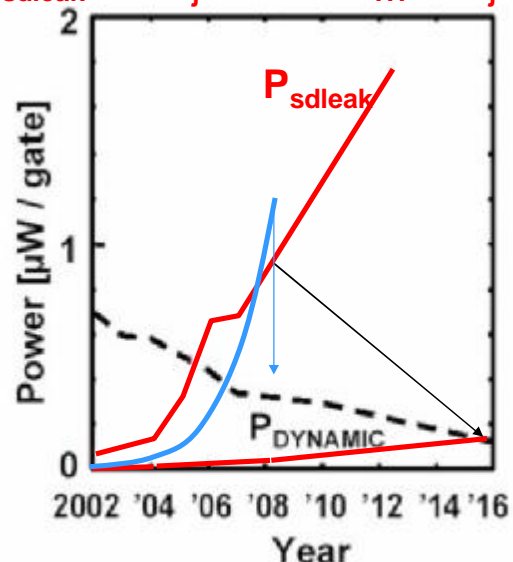
- **Leakage power starts to dominate**
 - ∅ larger gate delay than scaling for performance predicts
- **Voltage headroom shrinks**
 - ∅ makes A&RF guys deeply worried (sub 1 Volt circuits)
- **Interconnect claims the first role...**
 - ∅ challenging timing, power, synchronism, signal integrity
- **Increasing device variability**
 - ∅ jeopardizes predictability and yield, affects design process
- **New device architectures needed**
 - ∅ Impact layout style, need for DFM and IP library design



Leakage Power Issues



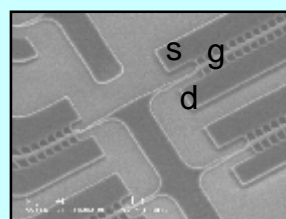
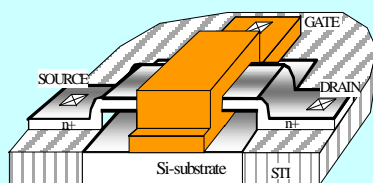
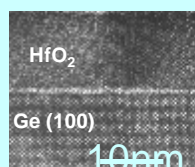
$$P_{sdleak} \sim W \cdot T_j^2 \exp(-qV_{TH}/nkT_j)$$



Hence "LateCMOS" from 65nm on ?

“LateCMOS” Device Architectures below 65nm

- Gate leakage: **Hi-k, metal (FUSI) gates (when?)**
- Mobility keeper: **Strained Si, (Ge...)**
- SD leakage: **Hi V_T , M- V_T , Hi- V_{DD}**
 \hat{U} dyn power, variability
- Ion boost: **Multi-Gate, FINFET, DGFDSOI**



2005

Emerald DGFDSOI
STM

Finfets, IMEC

2010

ISSCC

BUT Late CMOS diversity must percolate to system level...

Different Things to Different People...

For Microprocessor, Servers

@100 Watt active power, 20W leakage acceptable
 \Rightarrow low V_{TH} , high I_{on} , ABB to manage leakage yield

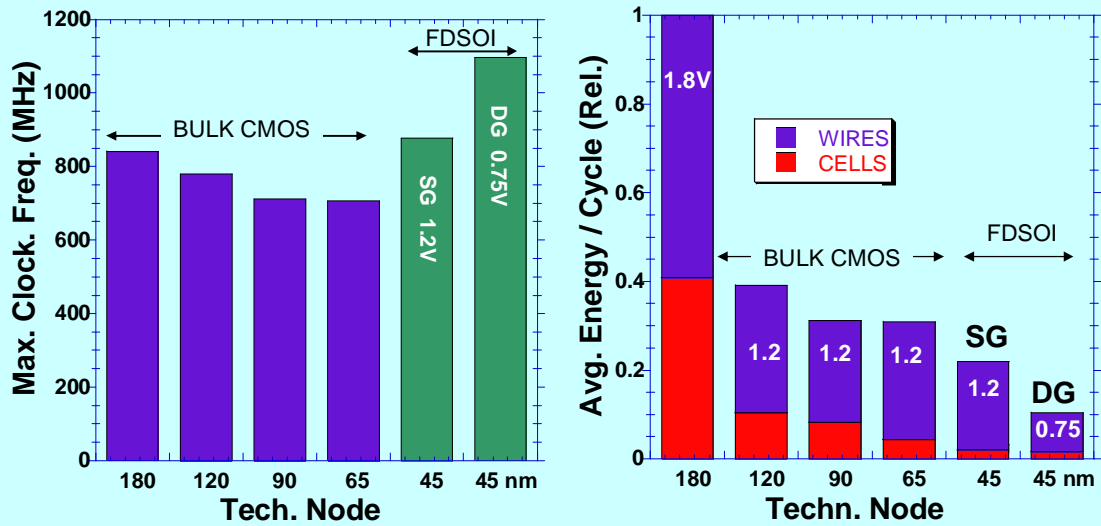
BUT for Aml...

Leakage mWatt nodes <10mW, μ Watt nodes <10 μ W
 \Rightarrow Scaling for low leakage at expense of gate delay
 \Rightarrow More GOPS from more transistors \hat{O} faster ones
 \Rightarrow Wires more important anyway...

Scaling for Low Leakage Aml

230 kgate Standard Cell Layout for 10pA/mm sd Leakage

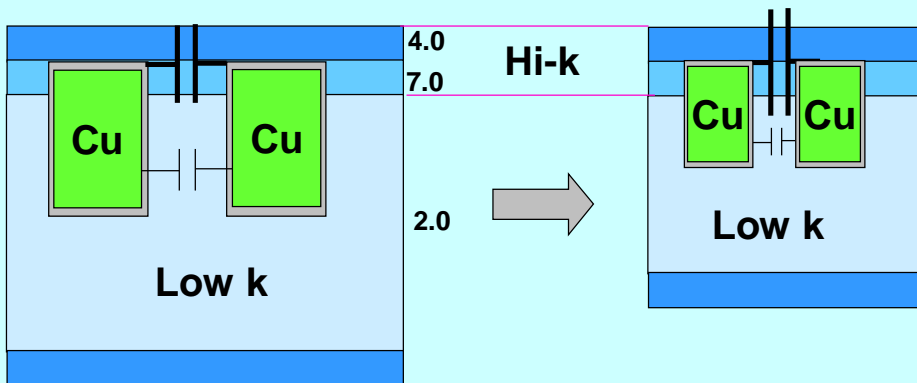
(Courtesy: Philips Techn. Modeling Group)



Below 65nm new device architectures needed
Local wiring power does not scale well

Nonlinear Scaling of Wiring

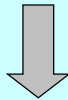
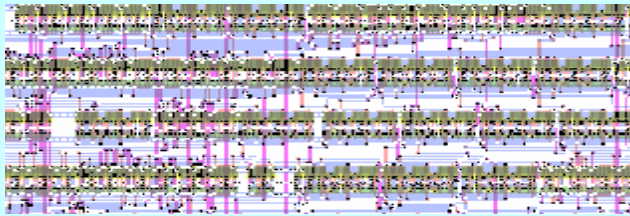
Hi-k Dielectric Barrier/Stop Layers don't scale
Absorb flux



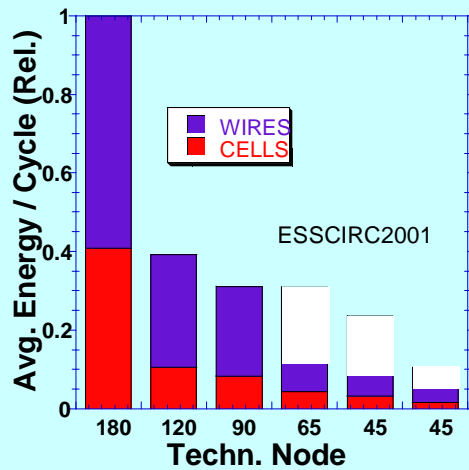
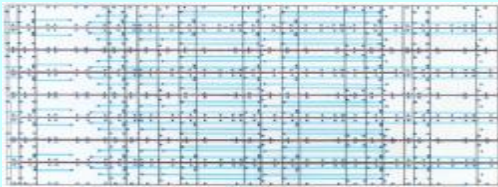
New dielectric barrier materials...
Keep wires further apart
Localize interconnect...

Back to the Future for Local Wiring?

From Standard Cell P&R

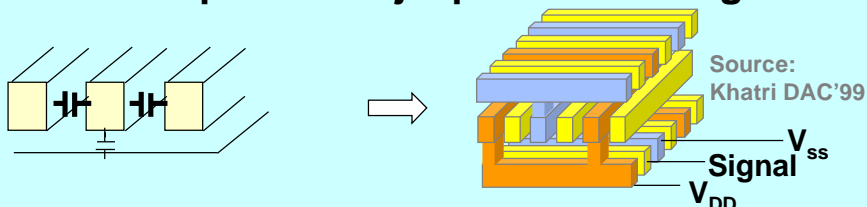


Back to tools for structured layout?

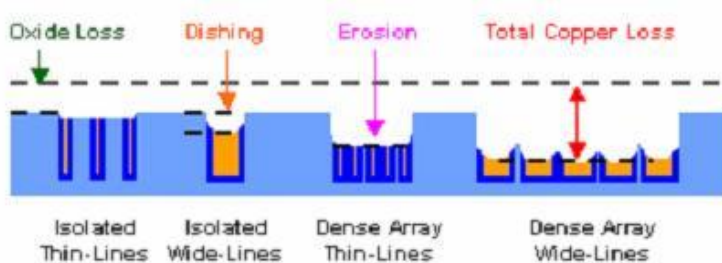


Global Interconnect Issues

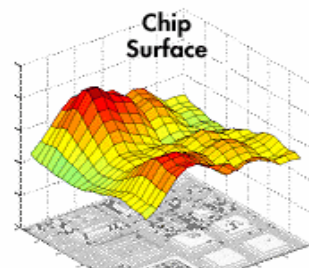
- Global Interconnect delay does not scale with logic
- Synchronous zones 45nm @ 500MHz < 2*2 mm²
- Interwire capacitance jeopardizes timing closure



- CMP Copper Thickness Variation up to 40%



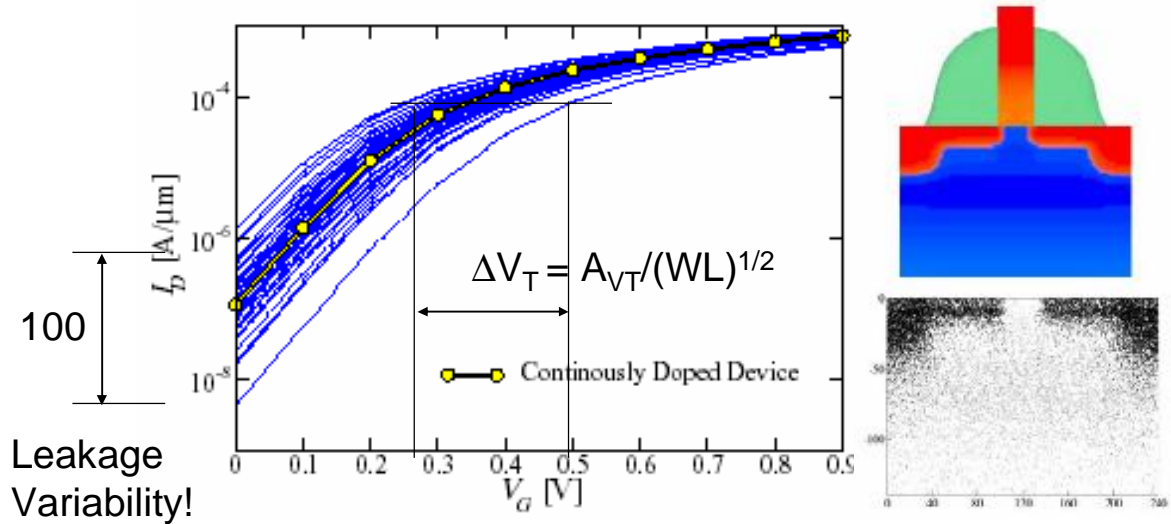
Source: Preagusus



Delay Variation



Atomistic Variability Shows Up...

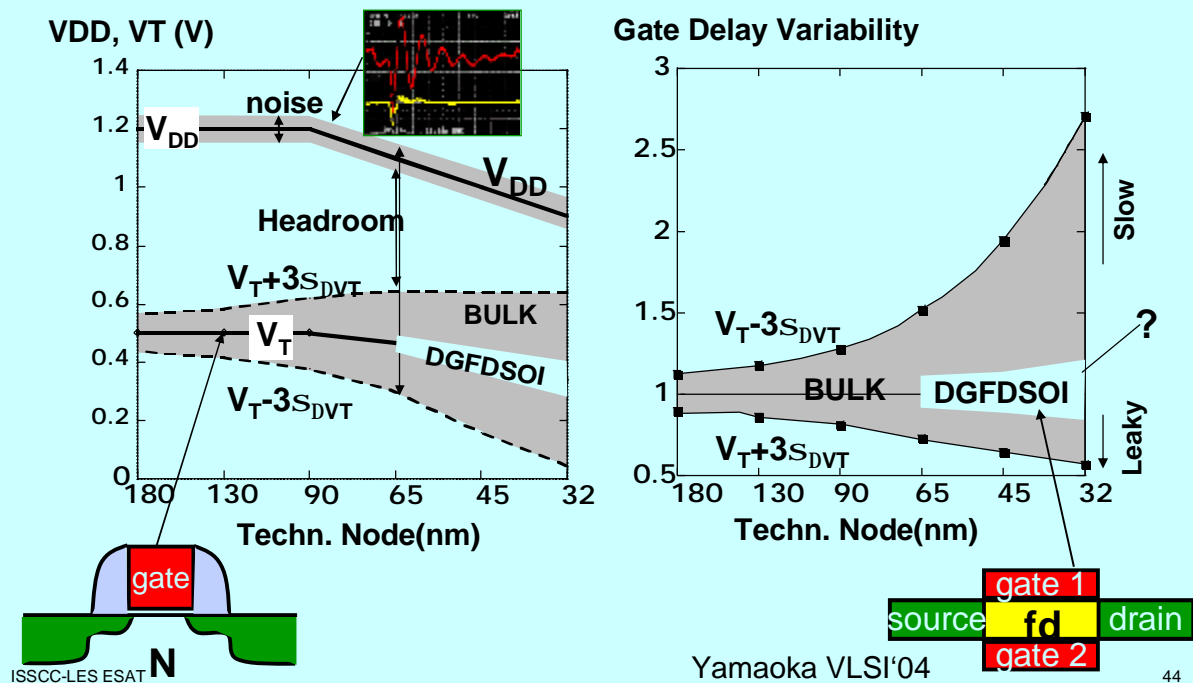


35nm Transistor
Asenov, ESSCIRC'04



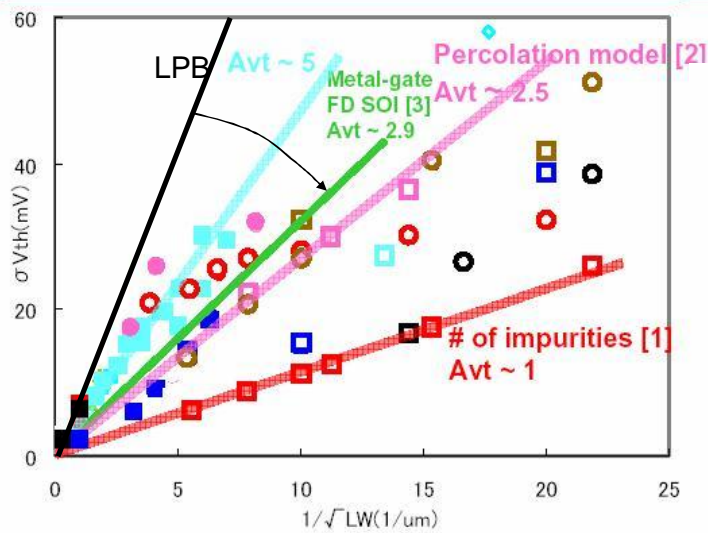
Variability Reduces Predictability

Low Standby Power CMOS; Min size transistors



Are "Late CMOS" devices better?

Investigate variability origin of FD MUGFETS



Avt (mV. μm) comparison

Technology	nmos	pmos
FD SOI	2.9	2.6
PD SOI	6.1	4.2
Low-Power Bulk	8.2	7.5

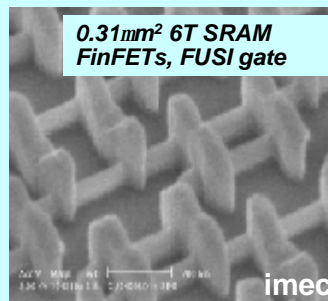
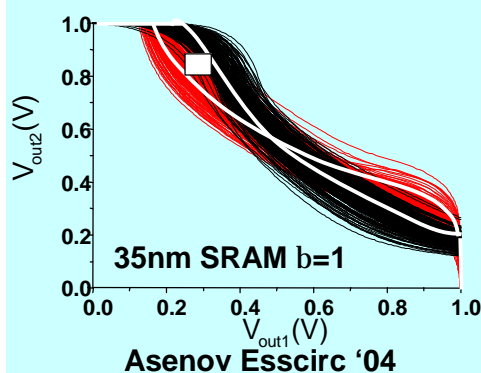
Vandooren IEDM 2003

[1] T. Mizuno et al. VLSI Tech. Symp, 1993. [2] A. J. Bhavnagarwala, IEEE JSSC, 2001. [3] A. Vandooren, IEDM, p. 975, 2003.

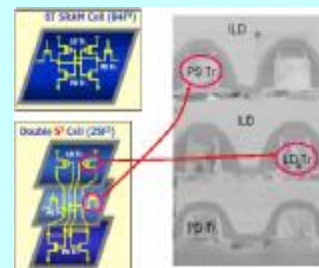
Source : T.Hiramoto VLSI 2005

Variability Impact

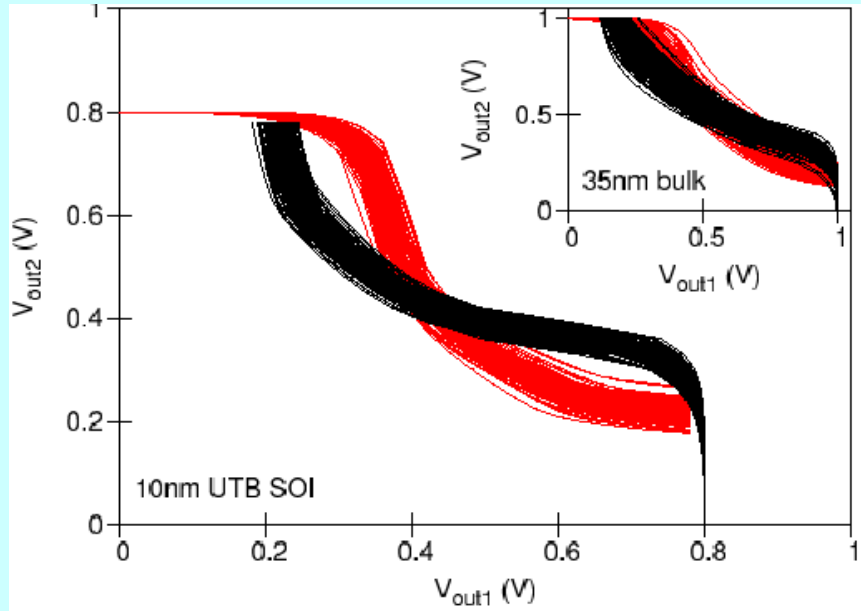
- Adds a new dimension to design space (P-A-T-Y)
- Statistical design, device sizing for yield crucial
- SRAM first victim...VDD scaling problematic
- Below 65nm new device architectures needed



**45nm FinFet
SRAM IEDM'04**



Conventional vs UTB MOSFETs



BISIMSOI is used for the UTB SOI MOSFET



But System Design must change...

Better than worst case design else lose advantage of scaling

Design reliable systems from uncertain components

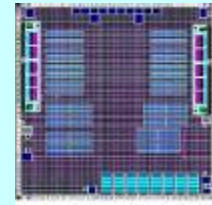
Allow errors to happen!





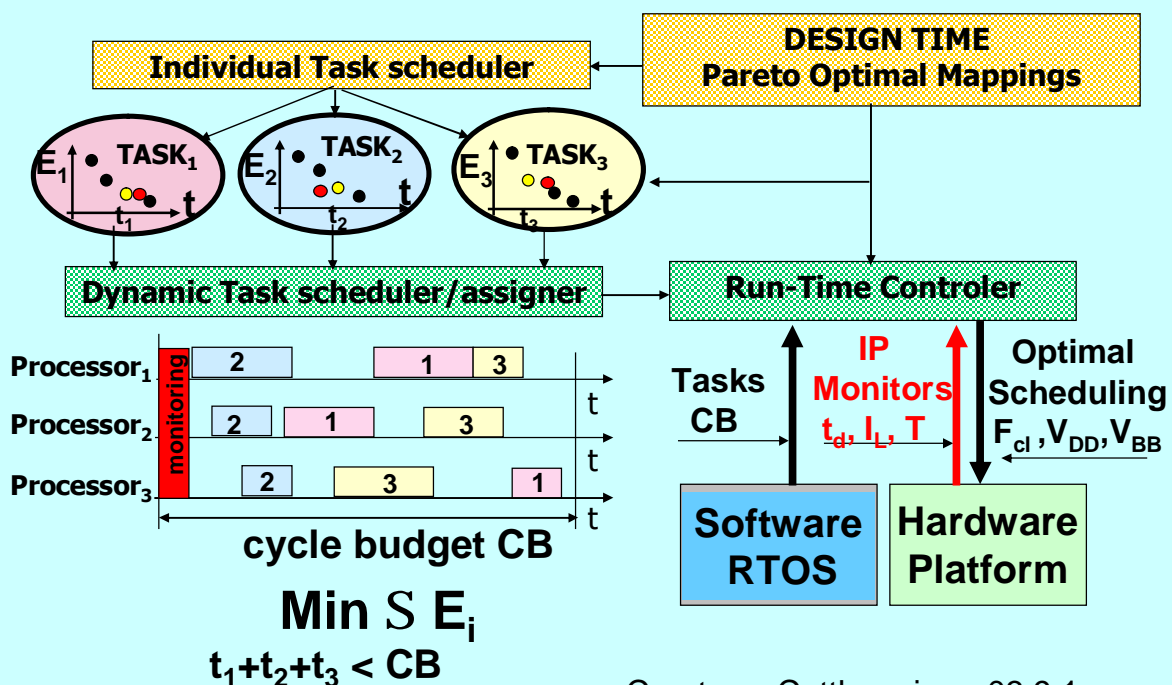
Ways to Cope...

1. Adapt V_{DD}, f_{cl}, I_L at run time to whatever mother nature provides and to dynamic nature of software tasks
2. Use tile based GALS architectures, networks-on-chip, and error correcting on-chip communication techniques
3. Back to regularity (Litho, context)



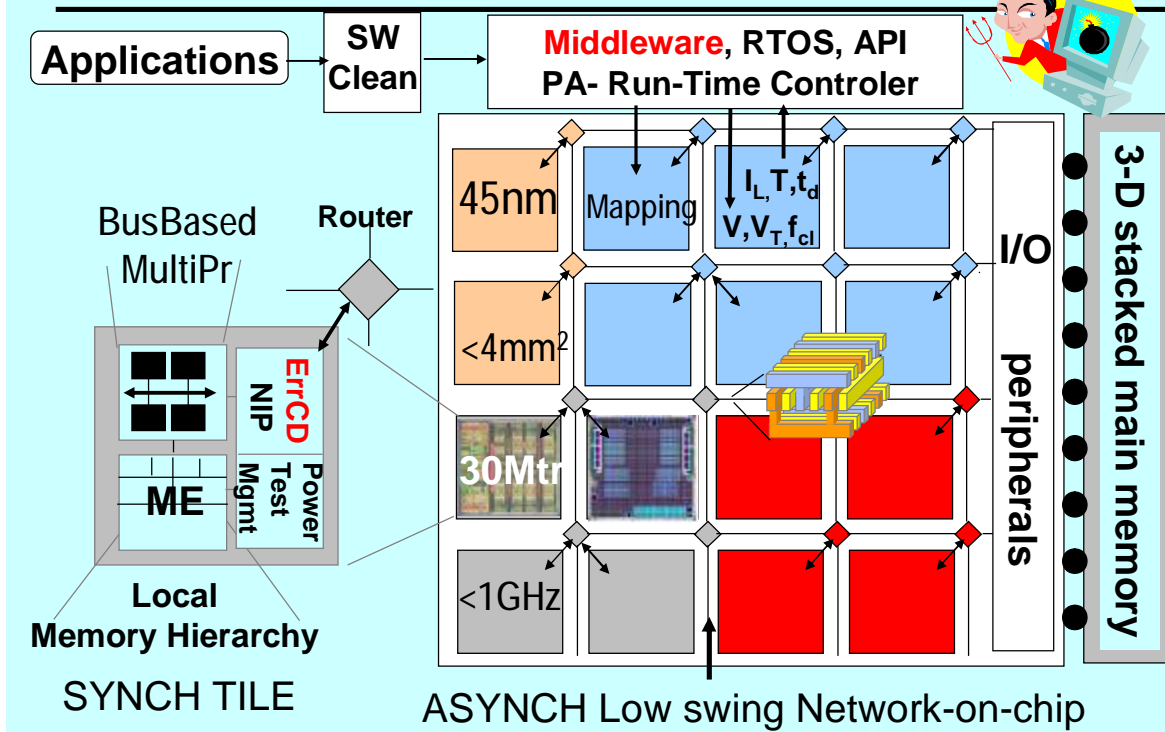
RapidChip

Run Time Adaptive Design

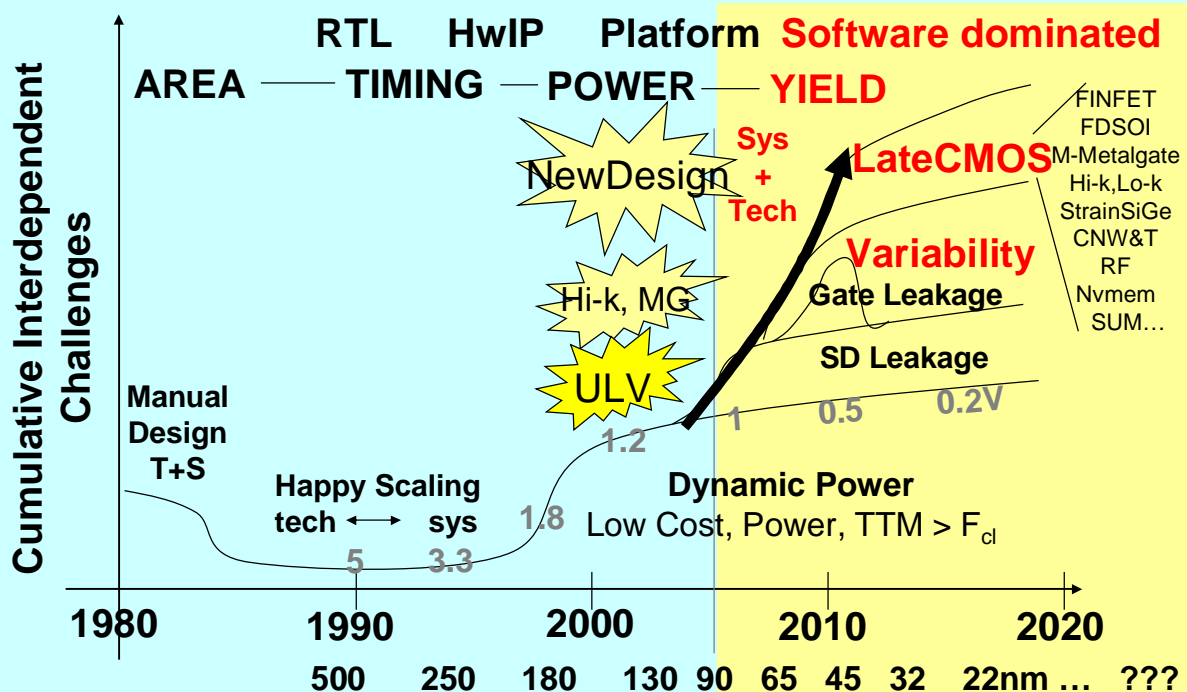


Courtesy: Cathoor isscc03 9.1

GALS Platform of Future + NOC



2005-2020 Not Business as Usual!



More Moore ...

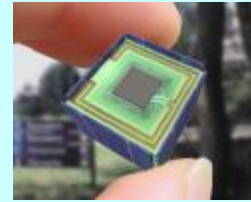
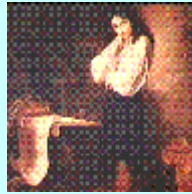
- **Roadmap beyond 65 nm fairly bumpy: the end of the SiO₂ era causes everything to change at the same time**
 - Ø Managing giga-complexity faced with CMOS diversification
 - Ø New design methods, tools and skills **urgently** needed
- **Must reconcile the hell of physics with heaven of Aml dreams**
 - Ø But easier to install equipment than to change people's mind
 - Ø New processing nodes 2 years, **new design tools take 10 years,**
- **Not for the faint of hearth**
 - Ø A few platforms by industry alliances and disciplined armies of engineers...
- **Need to share cost of advanced R&D**

Outline

- **The Ambient Intelligence Dream**
- **More Moore: Managing Giga-Complexity**
 - Ø of digital Watt and MilliWatt SoC's
- **More than Moore: Ultra-Creativity**
 - Ø for ultra-low-power and -cost interfaces
- **Conclusions**

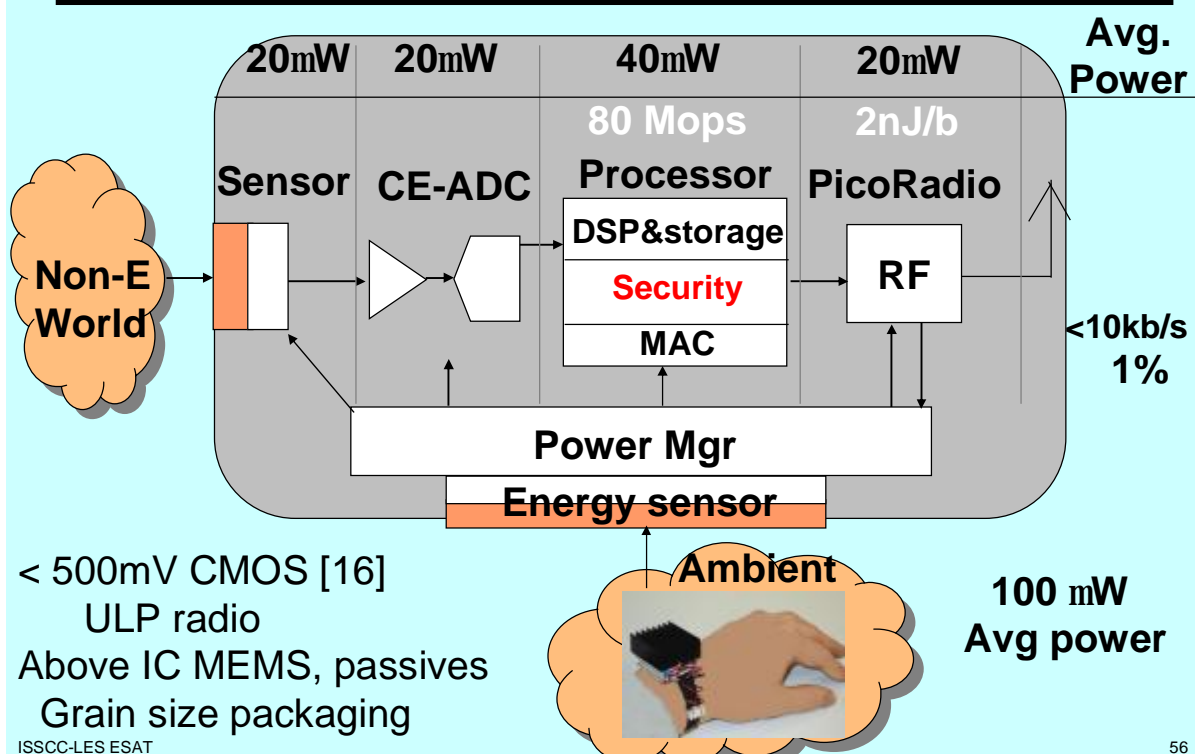
Interfacing to User and the Ambient

The art of ingenuity

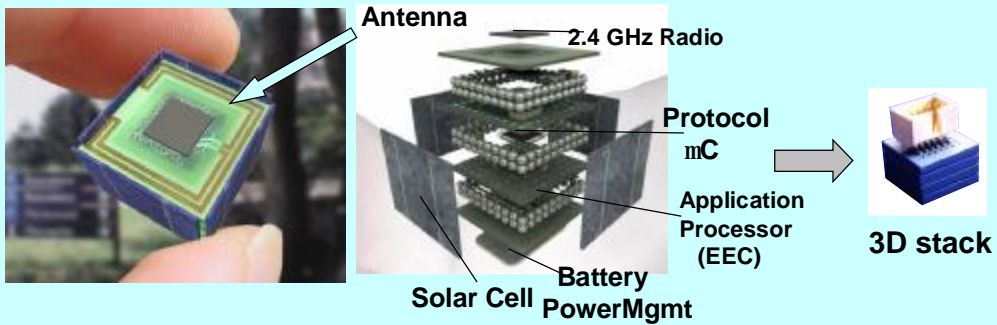


- **Get to the ultimate limits of**
 - Ø Miniaturization ($< 1\text{cm}^3$)
 - Ø Cost ($< 1\text{€}$)
 - Ø Power ($< 100\mu\text{W}$)
- **Design for utmost simplicity**
- **Interact with non-E world**
- **A micro-system node in ad-hoc network**

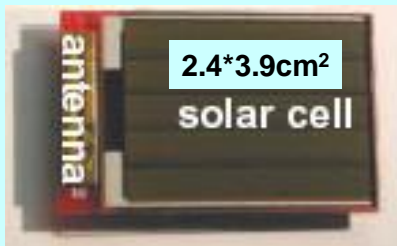
Sensor Node Challenges



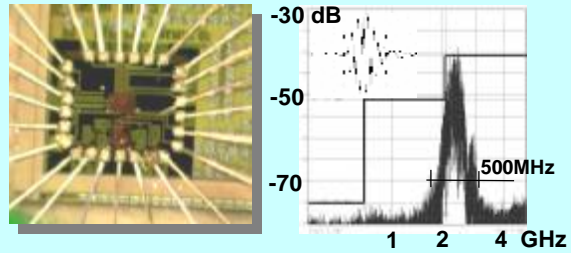
Sensor Radio's



(a) IMEC 1.4cm³, SiP Mote for EEC, ECG, 500mW@1%, 400b/sec

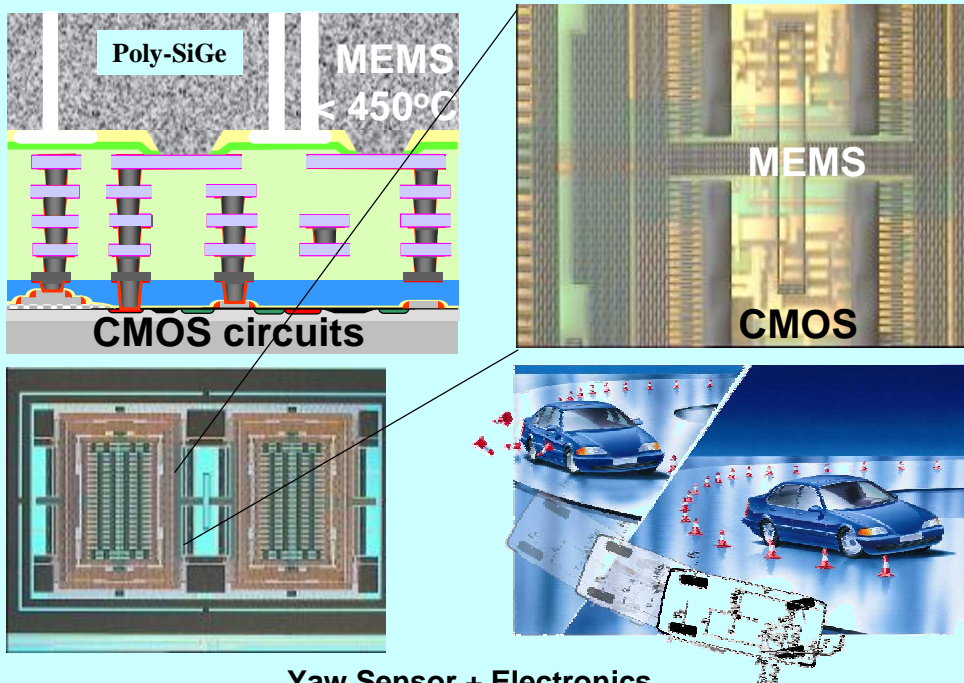


**(b) UCB PicoBeacon Tx [21.4]
1.9GHz, 400mW, 5kb/s**



**(c) IMEC 0.18m 0.25 mm² UWB Tx
0.5nJ/bit @ 10 kb/s**

Above IC MEMS and Hi-Q Passives [4,21]



**Yaw Sensor + Electronics
Courtesy: BOSCH-PHILIPS-IMEC**

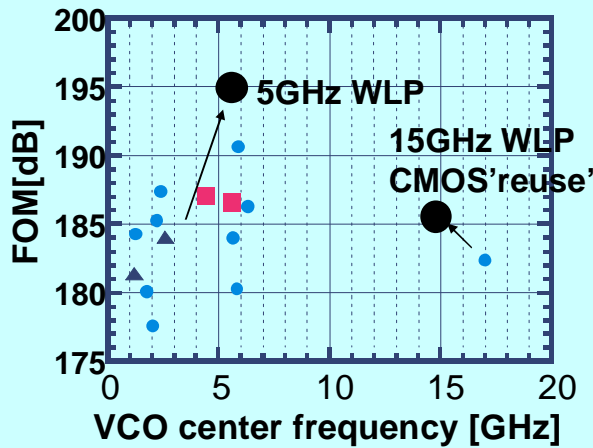
Above-IC VCOs on 90nm RF-CMOS

5 GHz VCO / 330 mW

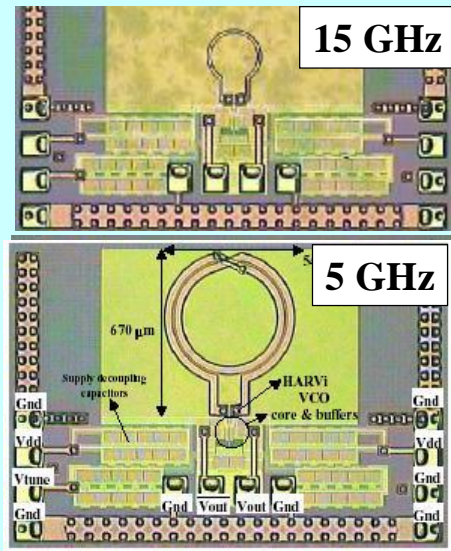
- 3nH WLP inductor
- $Q_{diff} \approx 40$ at 5 GHz
- -115dBc/Hz @ 1MHz
- 0.8 Volt

15 GHz VCO/2.5mW

- 0.6nH WLP inductor
- 105 dBc/Hz @ 1MHz

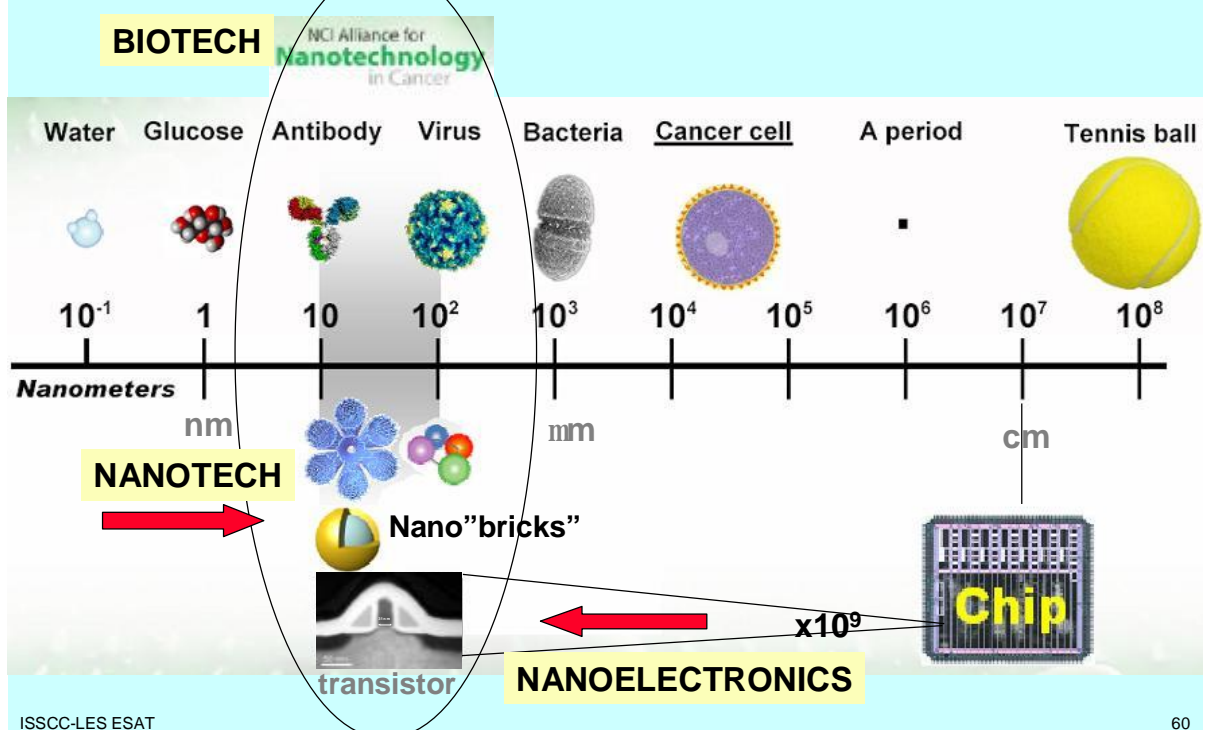


ISSCC-LES ESAT



Linten CICC04, ISSCC05[21] 59

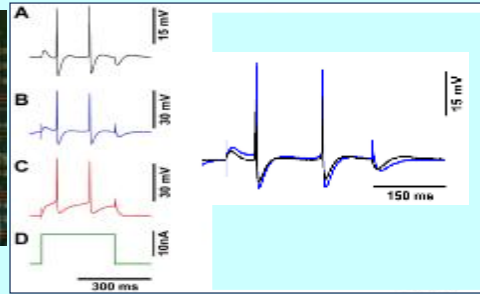
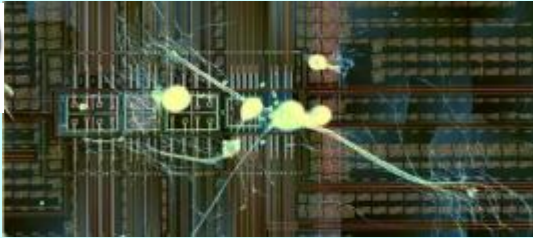
Convergence on the Nano-Scale: a unique opportunity



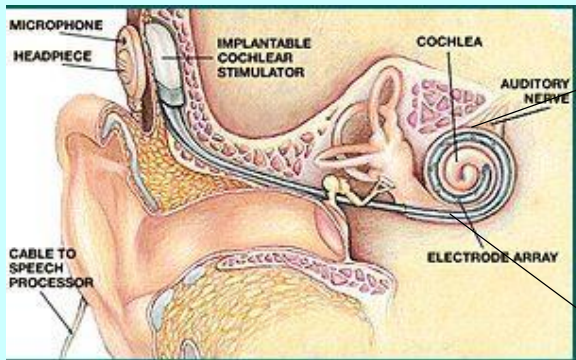
ISSCC-LES ESAT

60

Merging BIONics and ELECTRONics



NEUMIC *Aplysia California* culture on chip

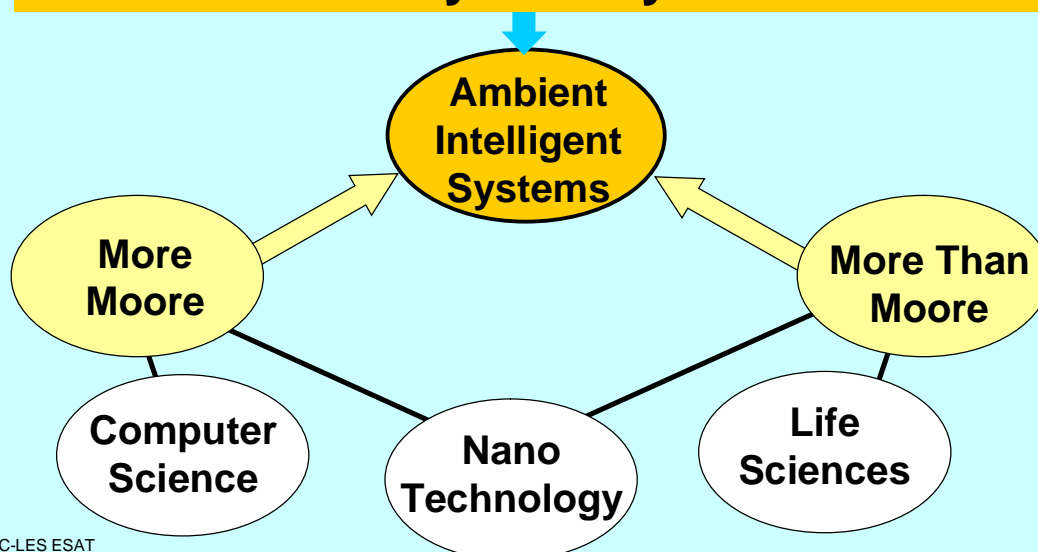


K. Wise, IEDM 2002

cochlear electrode array

Concluding Thought

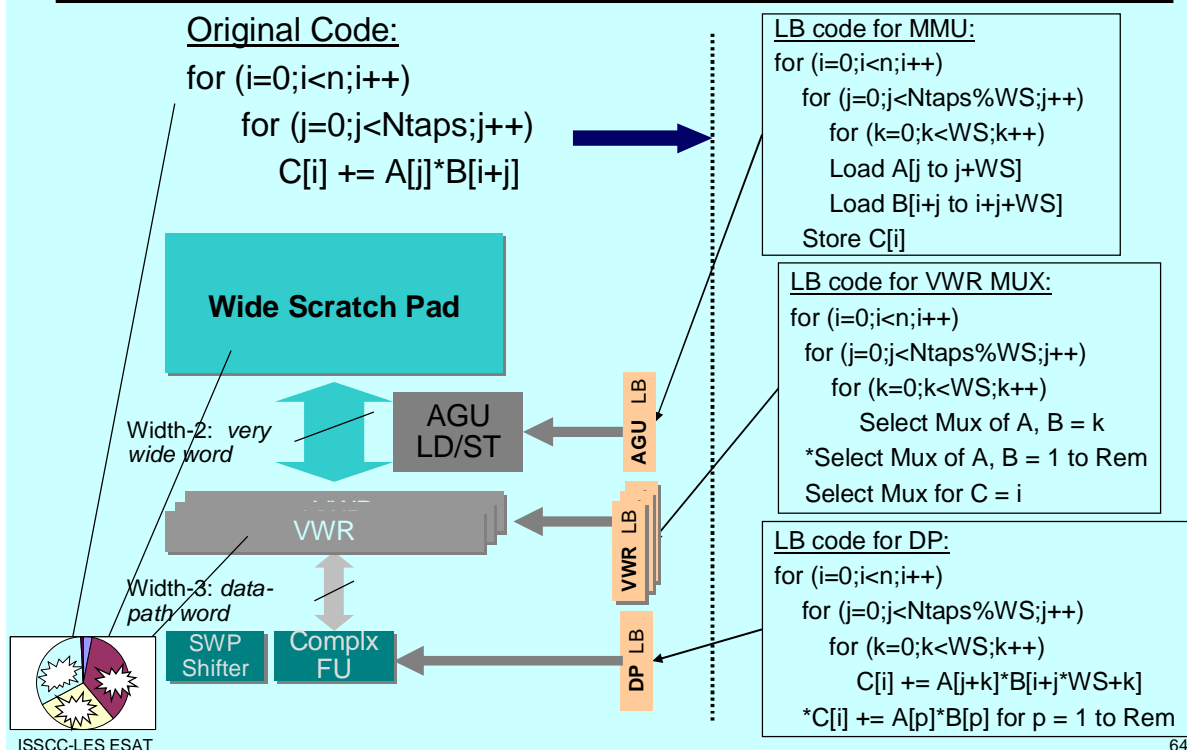
“Progress will be in the merger of previously unconnected techno-cultures driven by society needs”



Thanks to...

***E. Aarts, G. Beenker, T. Claasen, E. Persoon,
J. Huisken, P. Christie, W. Weber, R. Mertens,
R. Roovers, I. Young, W. Redman-White,
W. Geurts, C. Dallavalle, R. Zafalon, G. Borghs
F. Catthoor, K. Maex, J. Bormans, J Vounckx,
B. Gyselinckx, J. Van Ginderdeuren...
And so many others...***

Complexity is in the Compiler...



A Roadmap for Late-Silicon Age Design

