NEW NON-VOLATILE MEMORY SOLUTIONS: HOW THEY MAY (NOT) SERVE FUTURE SYSTEMS

LIESBET VAN DER PERRE

LUND CIRCUIT DESIGN WORKSHOP, OCT 3RD 2012

imec

MORE AND MORE MEMORY IS REQUIRED

On-chip memories today:

- > 50% total die-area and are responsible for more than
- > 40% total power consumption.
- Cache memory: 30% on-chip area in state-of-the-art microprocessors.

© IMEC 2012

EMBEDDED MEMORY MARKET SPACE DIVERSE APPLICATIONS

FPGAs

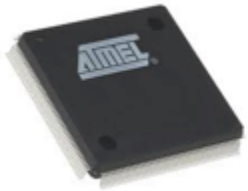
Mobile Application

Mobile Communication

Mobile Graphics

Desktop/servers

Micro-controllers



© IMEC 2012

NEW NON-VOLATILE MEMORY SOLUTIONS: HOW THEY MAY (NOT) SERVE FUTURE SOCS

1. Why move our cheese?

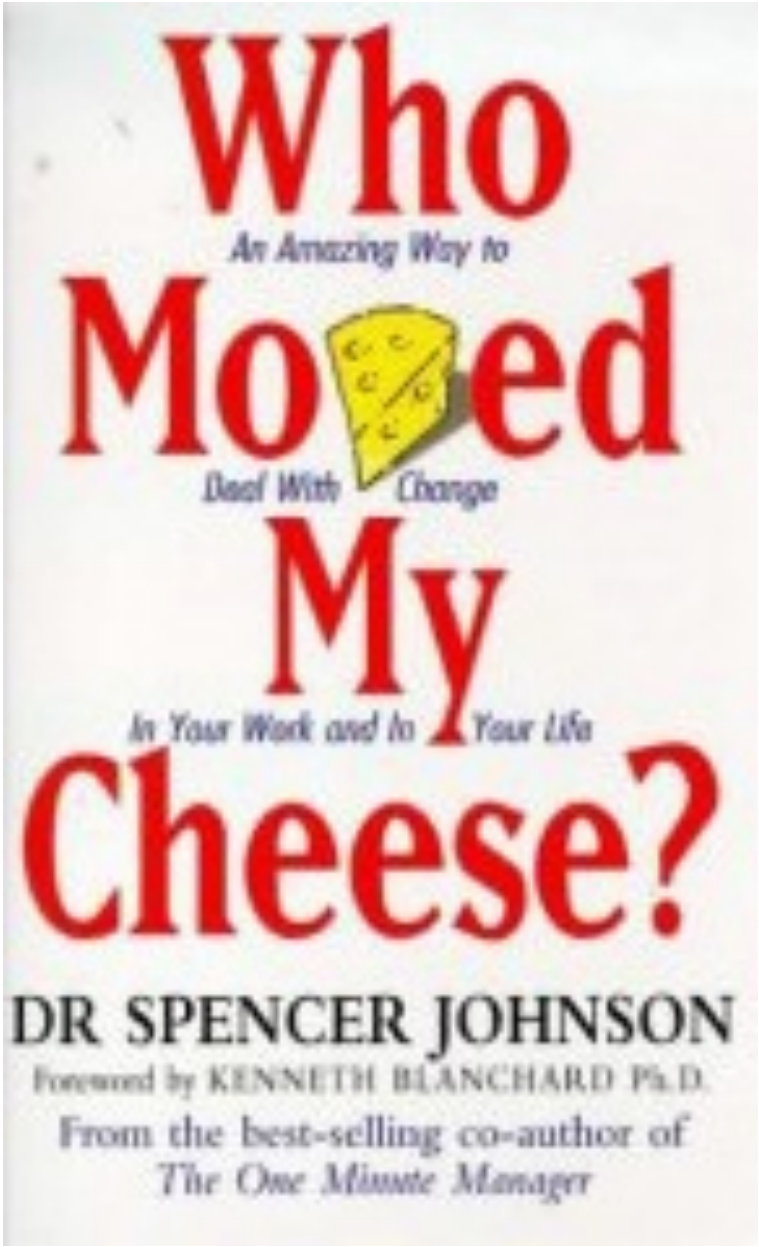
2. Memory technology: emerging non-volatile solutions

3. System challenges/solutions:
   - Write latency: scheduling to conceal the penalty
   - Endurance: dedicated error coding could help

© IMEC 2012

Warnings:

- Work (getting) in progress
- Wireless propagation background

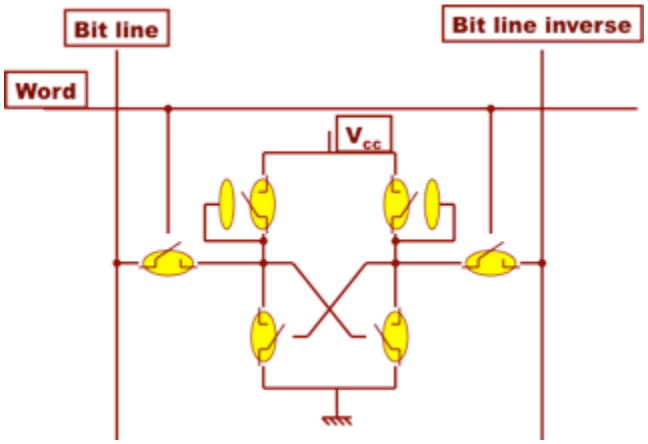


© IMEC 2012

SRAM: POPULAR SOLUTION TODAY

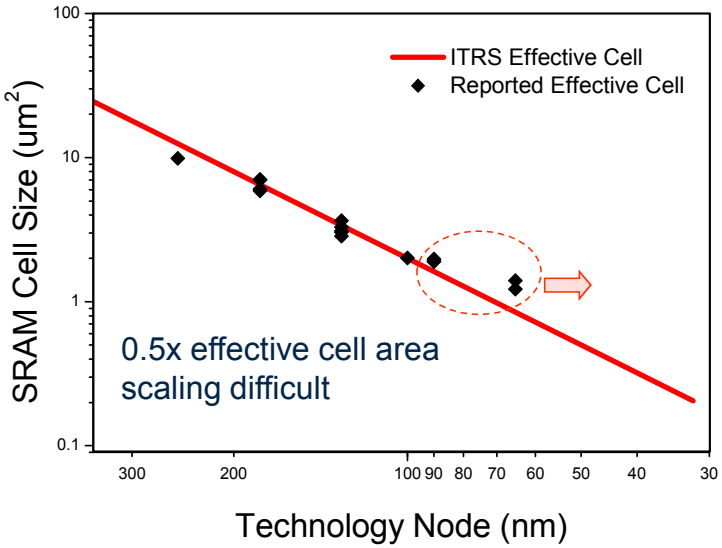
6T memory cell

- Fast
- Relatively low power



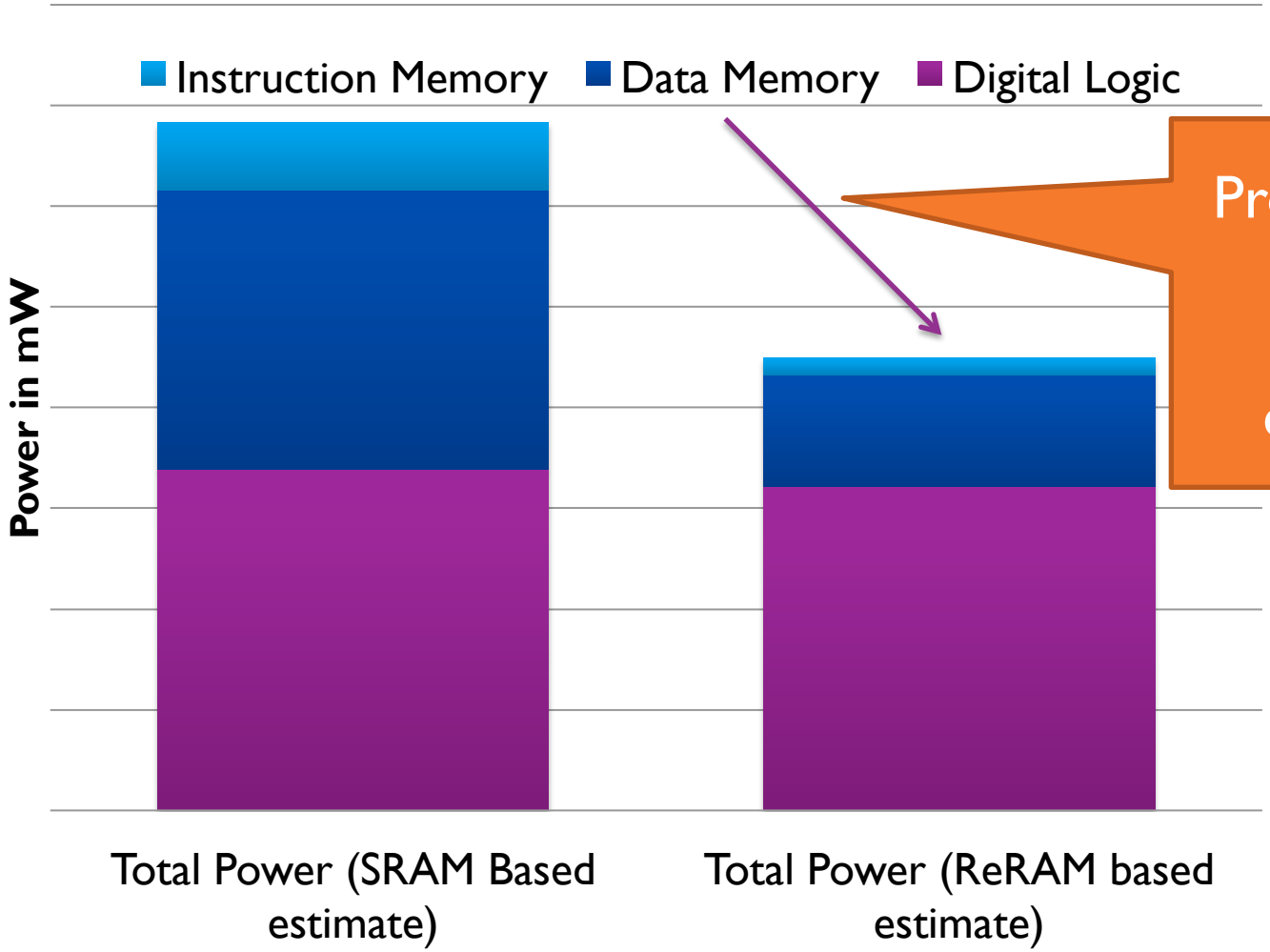
But: (how) will they scale?

- Increasing leakage
- Increasing variability



© IMEC 2012

EMBEDDED NVM IN RELEVANT PROMISE GREAT GAINS FOR SOC AREA (AND POWER?)

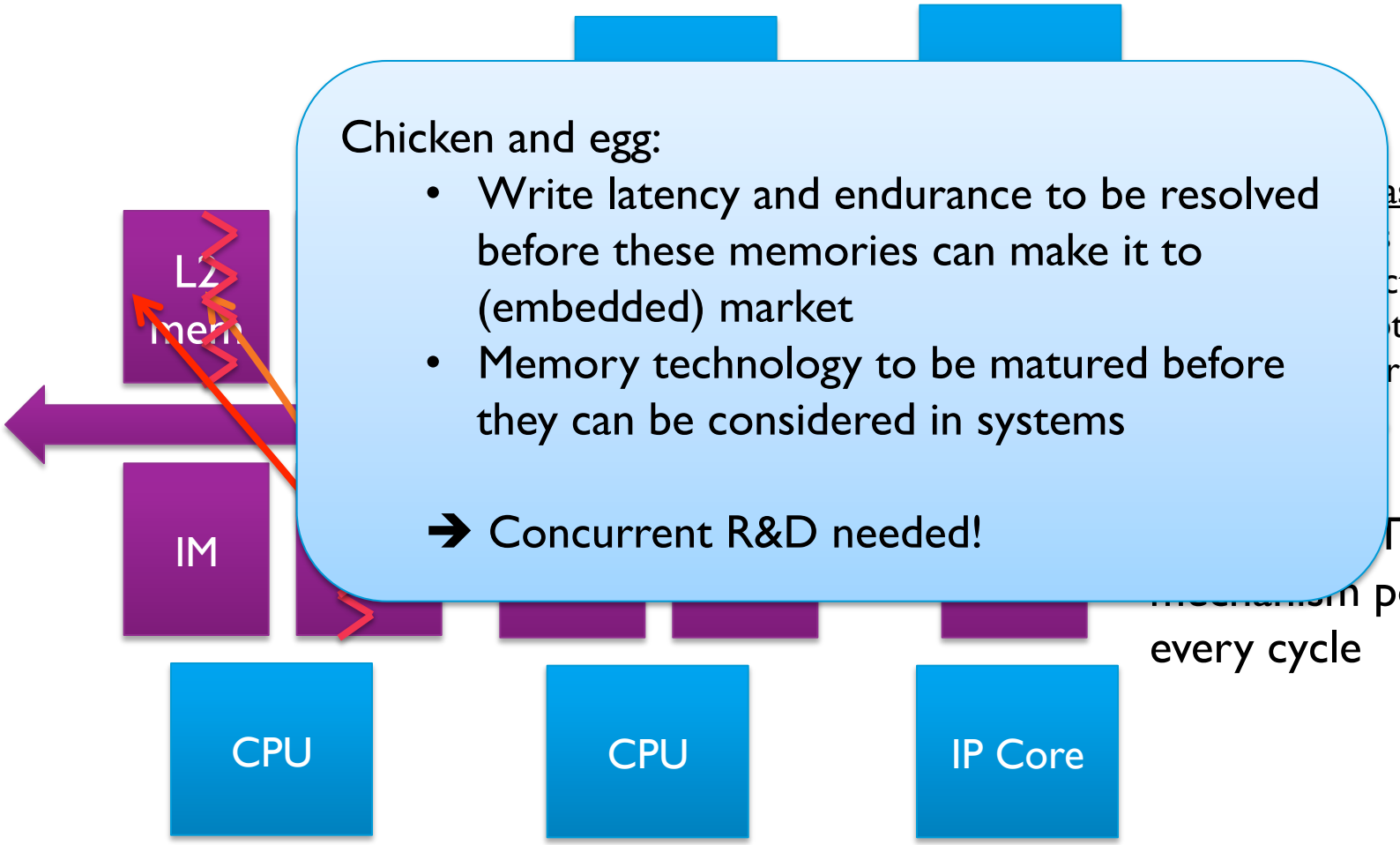


Projected potential power gains of 30-40% on the complete SoC?

Area gains can be expected of at least same or higher

CHALLENGE: OPERATE NEW MEMORIES WITHOUT PERFORMANCE PENALTY

Assumption Change: "Volatile embedded memory becomes non-volatile"



Chicken and egg:

- Write latency and endurance to be resolved before these memories can make it to (embedded) market
- Memory technology to be matured before they can be considered in systems

➔ Concurrent R&D needed!

ase: Transfer
to change →
cture needs to
t just plug and
r modifications)

Transfer
possible
every cycle

© IMEC 2012

NEW NON-VOLATILE MEMORY SOLUTIONS: HOW THEY MAY (NOT) SERVE FUTURE SOCS

1. Why move our cheese?

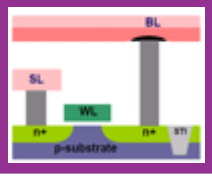
2. Memory technology: emerging non-volatile solutions

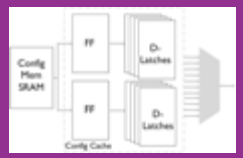
3. System challenges/solutions:
   - Write latency: scheduling to conceal the penalty
   - Endurance: dedicated error coding could help

© IMEC 2012

MEMORY @ IMEC: TECHNOLOGY & SYSTEMS WILL MEET IN THE MIDDLE?

What?	Memory cell/array (Emerging Memory)	System solutions (embedded/mobile applications)
Which value?	Memory stack and statistics	System solutions for endurance and latency
What does it take?	Material research Cell PDK	SoC architectures, ECC solution





Betting on two horses (RRAM & STT-MRAM)
They could both win races

© IMEC 2012

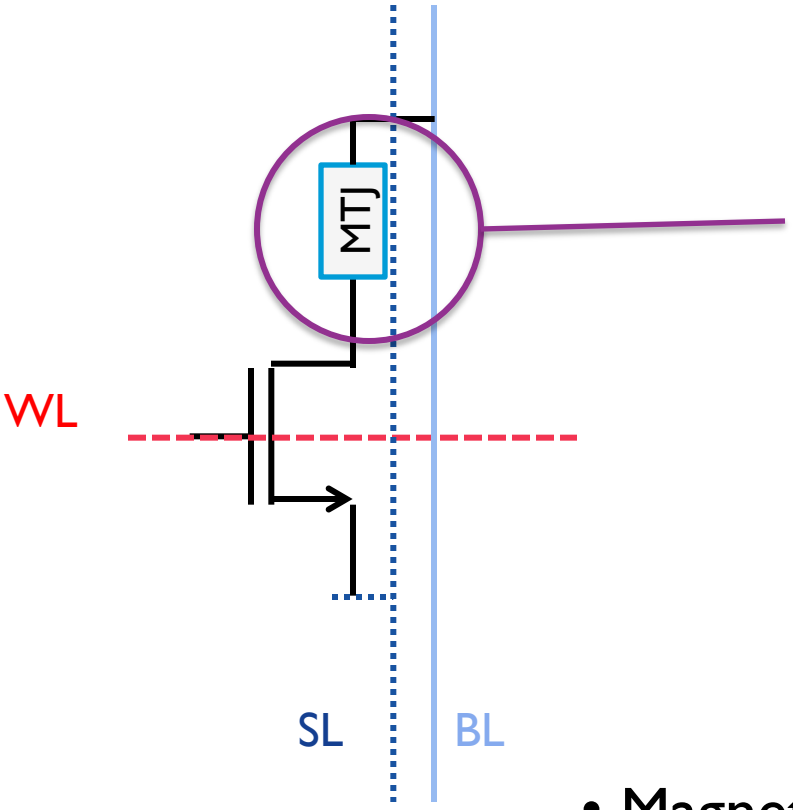
EMERGING NON-VOLATILE MEMORIES

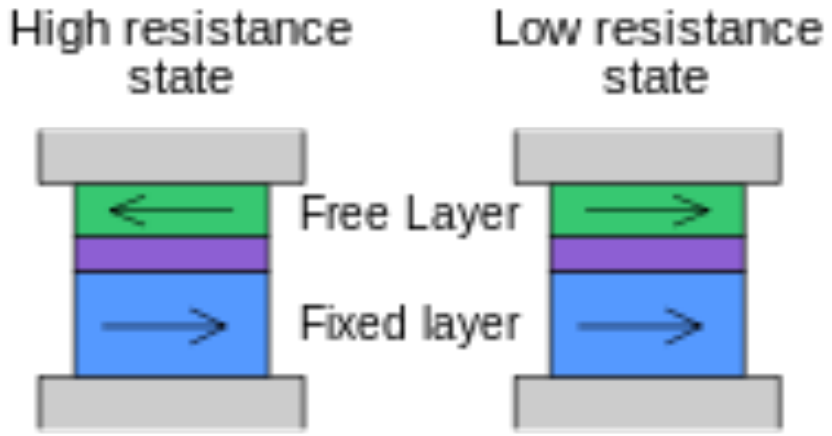
Memory	Flash	MRAM	PCRAM	FeRAM	RRAM	CBRAM
Write current	5 μ A	0.87mA	1.2mA	N/A	25 μ A	10 μ A
Write time	200ms	~ 10ns	~ 300ns	10ns	5ns	50ns
Access time	15ns	8ns	12ns	8ns	8.5ns	50ns
Endurance	~10^5	10^{15}	10^6	N/A	>10^{10}	10^6
Cell area	Small	Medium	Medium	Medium	Medium	Medium
R Ratio	> 10	2	> 10	N/A	> 10	>100

Phase Change RAM (PCRAM),
Ferro-Electric RAM (FeRAM)
Conductive Bridging RAM (CBRAM)

© IMEC 2012

SPIN-TRANSFER-TORQUE (STT-)MRAM CELL

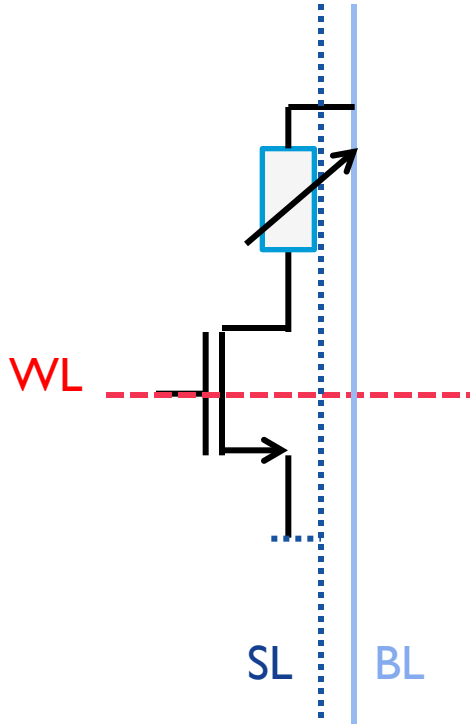




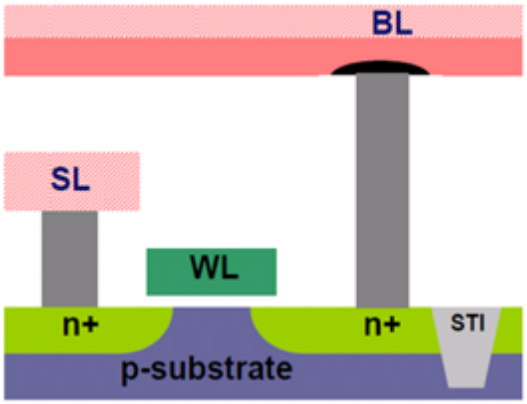
- Magnetic Tunnel Junctions used
- Promising write time, area and endurance
- Limited R ratio, cost?

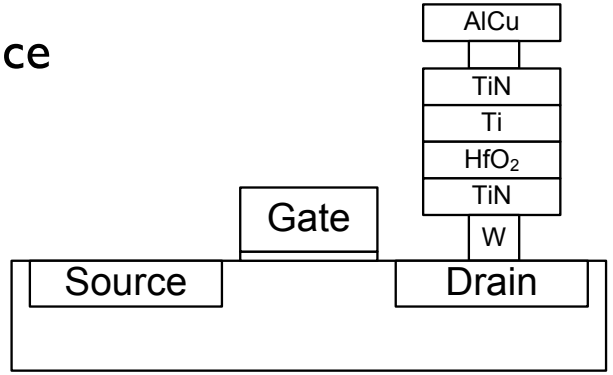
© IMEC 2012

RRAM: CMOS COMPATIBLE & DENSE



- Memristor used as storage element
- Retains the last resistance that it had when the current stopped
- Based on the switching mechanism
- Low cost
- High R ratio
- Low endurance



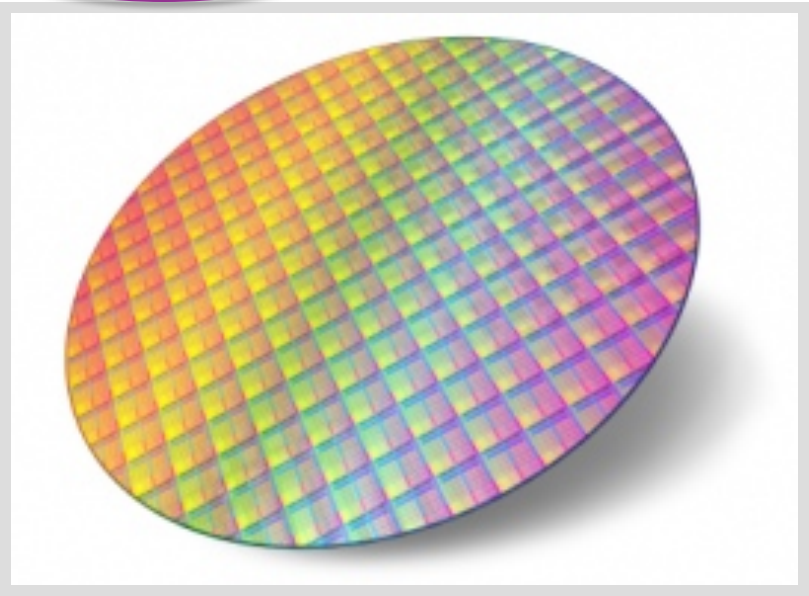


© IMEC 2012

Technology News

Taiwan embeds ReRAM in 28-nm logic process

September 26, 2012 // Peter Clarke



The ability to embed resistive RAM (ReRAM) into mainstream logic process technology could have major implications for the design of system-on-chip, and foundry chip maker Taiwan Semiconductor Manufacturing Co. is keeping tabs on the emerging capability.

One of the more intriguing papers listed in the advance program for the 2012 International Electron Devices Meeting in December is authored by a research team from National Tsing-Hua University (Hsinchu, Taiwan) that is also affiliated with TSMC.

The abstract to the paper, titled "High-K metal gate contact RRAM (CRRAM) in pure 28-nm CMOS logic process," states that a contact RRAM (CRRAM) cell has been realized in a HKMG 28-nm CMOS logic process without the use of any additional masking or process steps. This was done using a 35-nm by 35-nm contact hole.

© IMEC 2012

‘The amount of memory included on system chips is growing and becoming increasingly responsible for much of the power consumption. The facility to hold data in dense non-volatile memory on a SoC--rather than in power consuming SRAM for even a few processor cycles--could drive power savings in leading- edge manufacturing.’

© IMEC 2012

MANY MEMORIES IN EACH TARGET DOMAIN WITH DIFFERENT SPECIFICATIONS

Market	Memories	Technology Flavor used	Core Freq Range
FPGA	BRAMs	HPM or HPL	200MHz-500MHz
Mobile Communication	L1, L2 Instruction L1, L2 Data	HPM	400MHz
Mobile Application	L1, L2 Instruction L1, L2 Data	HPM	1.2-1.5GHz
Mobile Graphics	L1, L2 Instruction L1, L2 Data	HPM	1.2-1.5GHz
Server/Desktop	L1, L2, L3 Instruction L1, L2, L3 Data	HP	1.5-3.6GHz
Micro-controllers	L1, L2 Instruction L1, L2 Data	HPL/LP	<100MHz

© IMEC 2012

VIEW OF REQUIREMENT OF DIFFERENT MEMORIES FROM DIFFERENT MARKET SEGMENTS

Memory	Area %	Speed (MHz)	Power % (incl Leak)	R freq	W freq
FPGA BRAM	>50%?	200-400	??	=	=
Mobile Comm. D, I	>25%, >20%	200-600	35%, 15%	↑,↑	↑,↓
Mobile Appln D,I	>20%, >20%	400-1200	30%, 15%	↑,↑	↑, =
Graphics D+I	>40%	400-600	30%, 20%	↑,↑	↑, =
Server D+I	±50%	1000-3200	15%, 15%	↑,↑	↑, =
MicroController D/I	>15%, >15%	50-100	20%	=, ↑	=,↓

Focus of our research: One of the most demanding

© IMEC 2012

NEW NON-VOLATILE MEMORY SOLUTIONS: HOW THEY MAY (NOT) SERVE FUTURE SOCS

1. Why move our cheese?

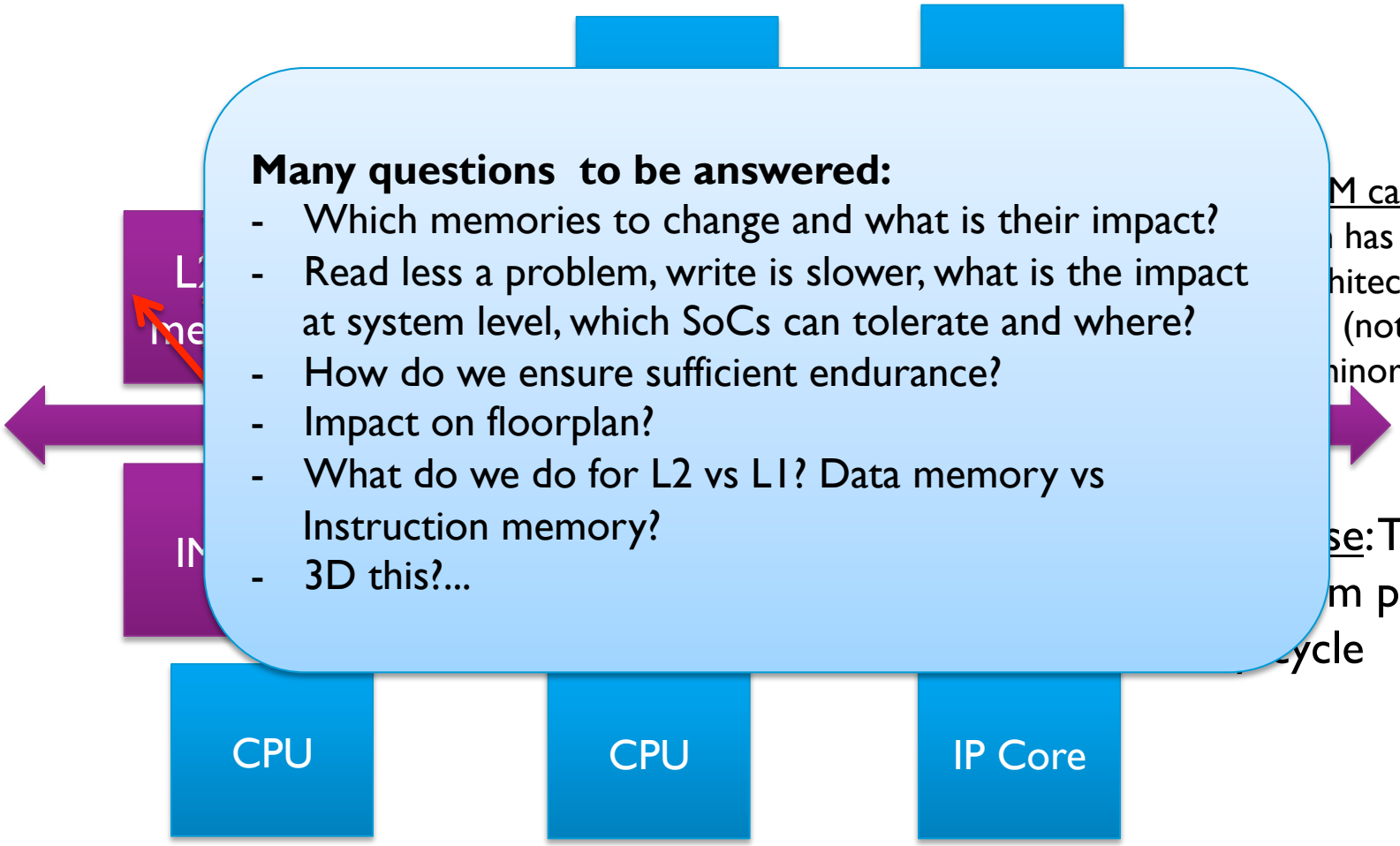
2. Memory technology: emerging non-volatile solutions

3. System challenges/solutions:
   - Write latency: scheduling to conceal the penalty
   - Endurance: dedicated error coding could help

© IMEC 2012

CHALLENGE: SYSTEM PERFORMANCE MAY GO DOWN (A LOT!)

Assumption Change: "Volatile embedded memory becomes non-volatile"



M case: Transfer
has to change →
hitecture needs to
(not just plug and
inor modifications)

se: Transfer
m possible
ycle

© IMEC 2012

RERAM VS SRAM: ENERGY GAINS COULD BE SUBSTANTIAL

Read Energy Consumption	64 Kbit SRAM		64 Kbit ReRAM	
	32 bit word read access	512 bit wide word read access	32 bit word read access	512 bit wide word read access
Read energy per access	4.39 pJ	32.41 pJ	0.74 pJ	2.20 pJ
Read energy per accessed bit	0.13 pJ	0.063 pJ	0.023 pJ	0.004 pJ

© IMEC 2012

PRELIMINARY CONCLUSIONS?

- Be aware potential performance degradation!
- Adapt memory architecture and scheduling for area and energy purposes
- The case considered: most challenging, many easier/ more gains

© IMEC 2012

NEW NON-VOLATILE MEMORY SOLUTIONS: HOW THEY MAY (NOT) SERVE FUTURE SOCS

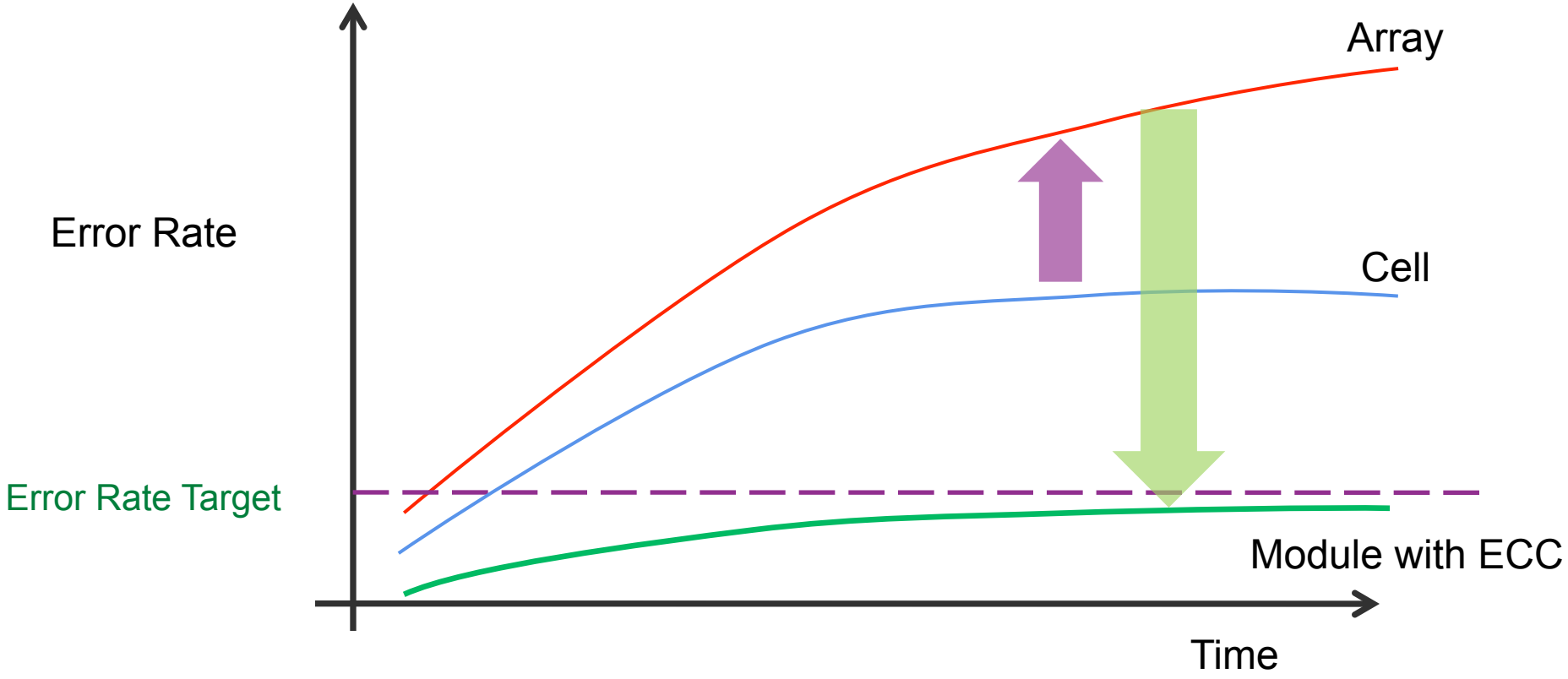
1. Why move our cheese?

2. Memory technology: emerging non-volatile solutions

3. System challenges/solutions:
   - Write latency: scheduling to conceal the penalty
   - Endurance: dedicated error coding could help

© IMEC 2012

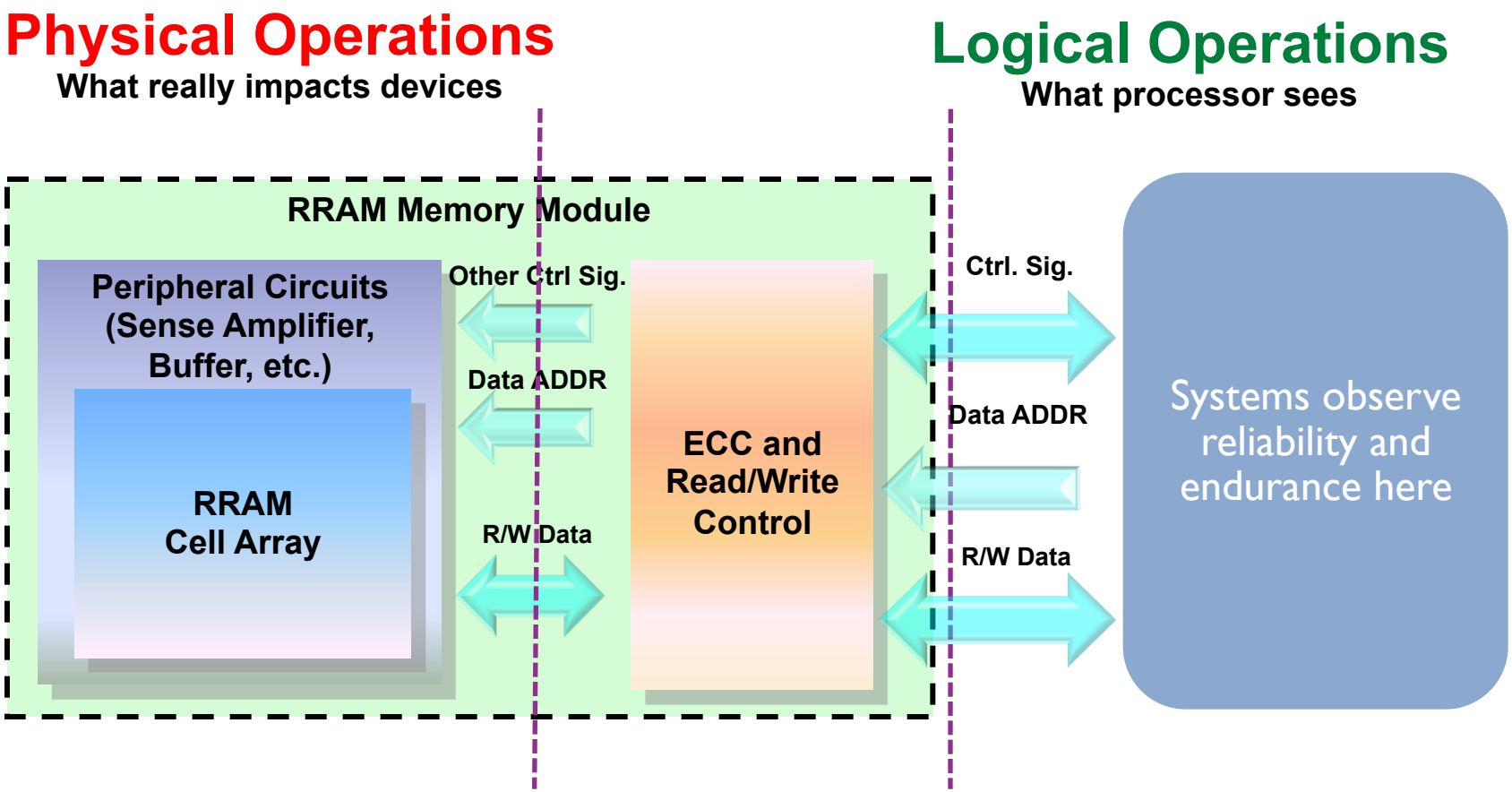
ECC SHOULD IMPROVE THE ENDURANCE OF RRAM MEMORIES



* Redundancy cells and encoding/decoding cost/power as overhead

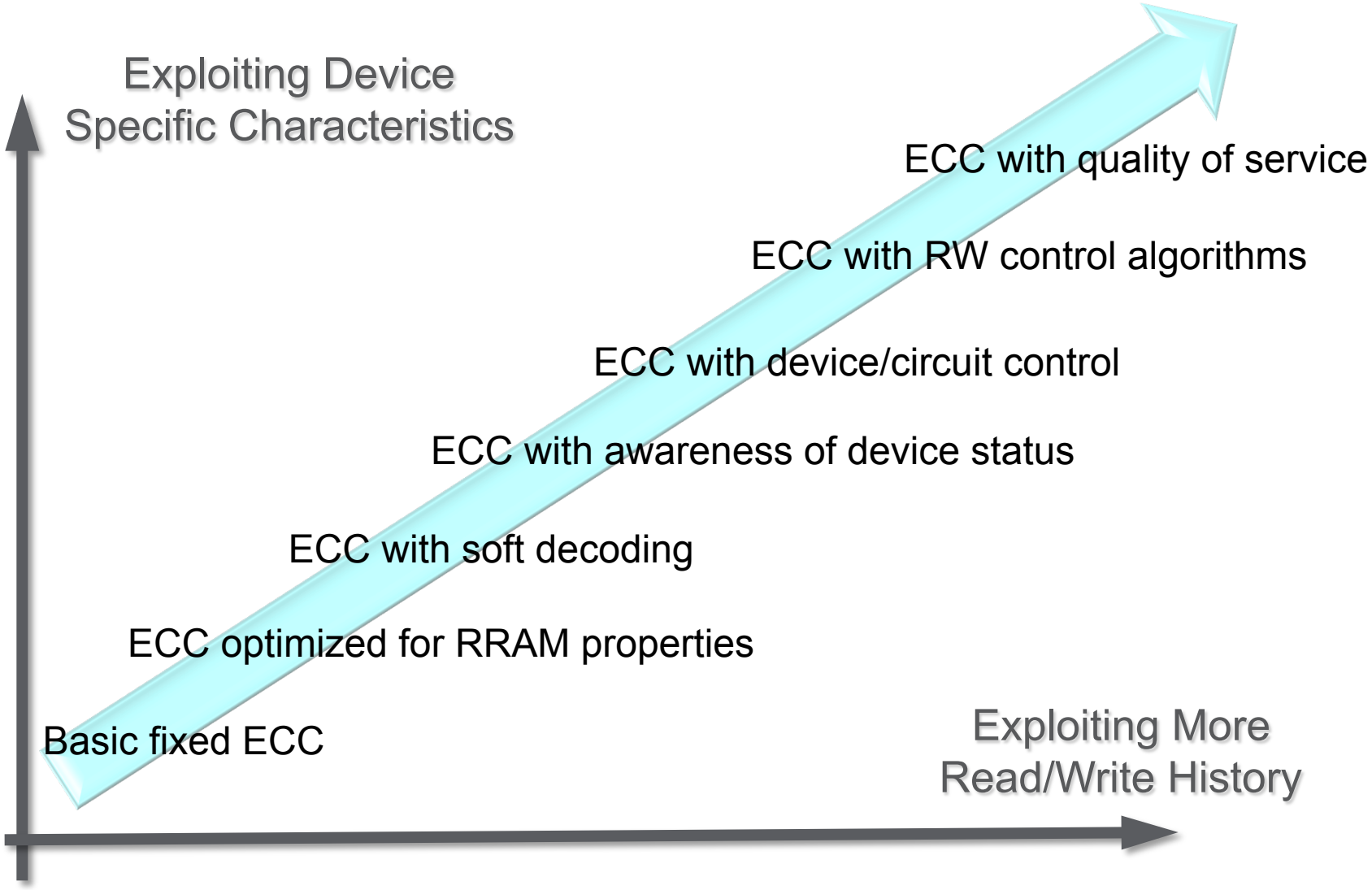
© IMEC 2012

ECC FOR RRAM: CRUCIAL FOR EMBEDDED AND STAND-ALONE

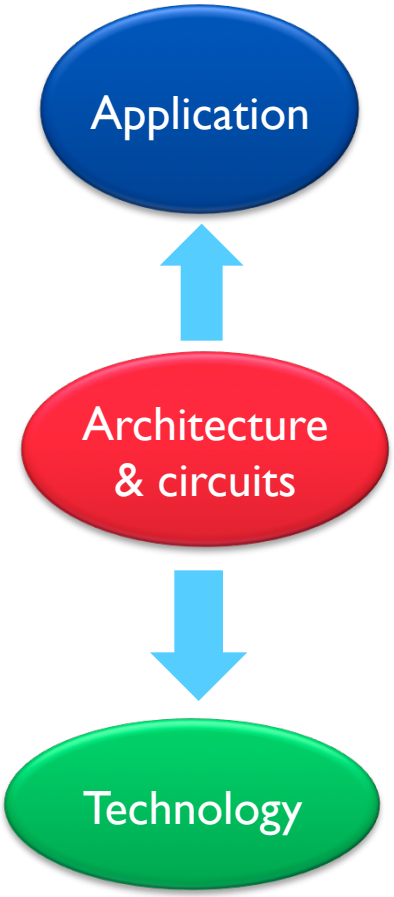


© IMEC 2012

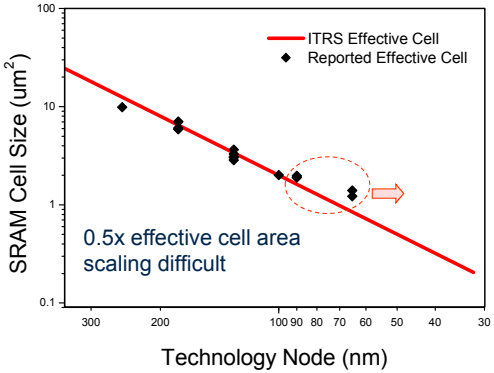
ADVANCED DEVICE DRIVEN ECC FOR RRAM: MANY IDEAS FOR RESEARCH



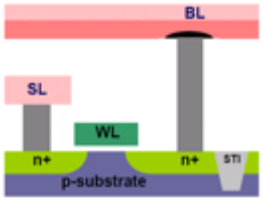
imec
© IMEC 2012



- SRAM scaling increasingly problematic



- New non volatile memories: attractive for (mobile) SoCs



- System & technology: gap to bridge

- And: non-volatile feature could be exploited!

Never give your old cheese away in anticipation of finding new cheese

© IM

THANK YOU

?

imec