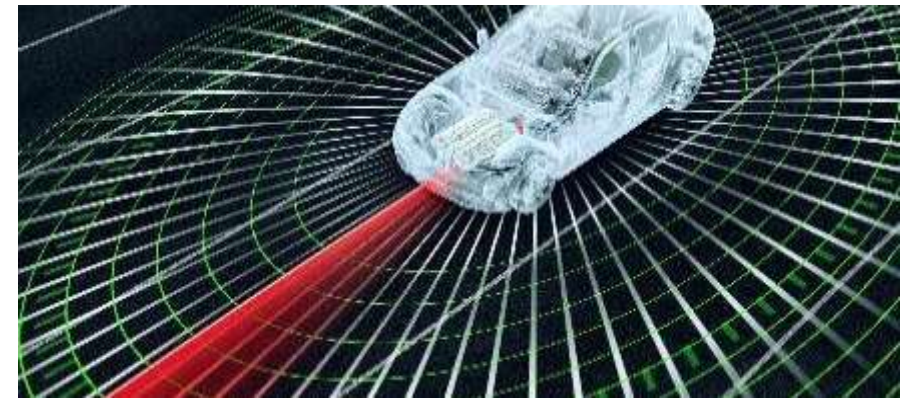
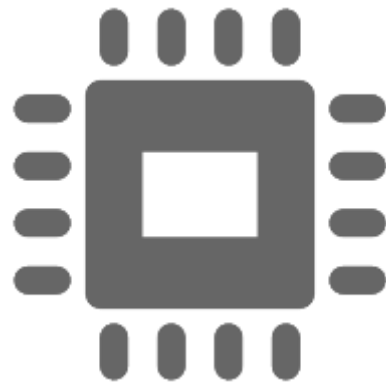


FROM RESEARCH TO INDUSTRY

cea tech



CHIP AGEING

ARCHI 2017 | Olivier HERON | March 9th, 2017



PART I: CONTEXT

A BIG TREND (HYPE?): CPS AND IOT

HiPEAC
COMPILATION ARCHITECTURE



HiPEAC Vision 2017

HIGH PERFORMANCE AND EMBEDDED ARCHITECTURE AND COMPILATION

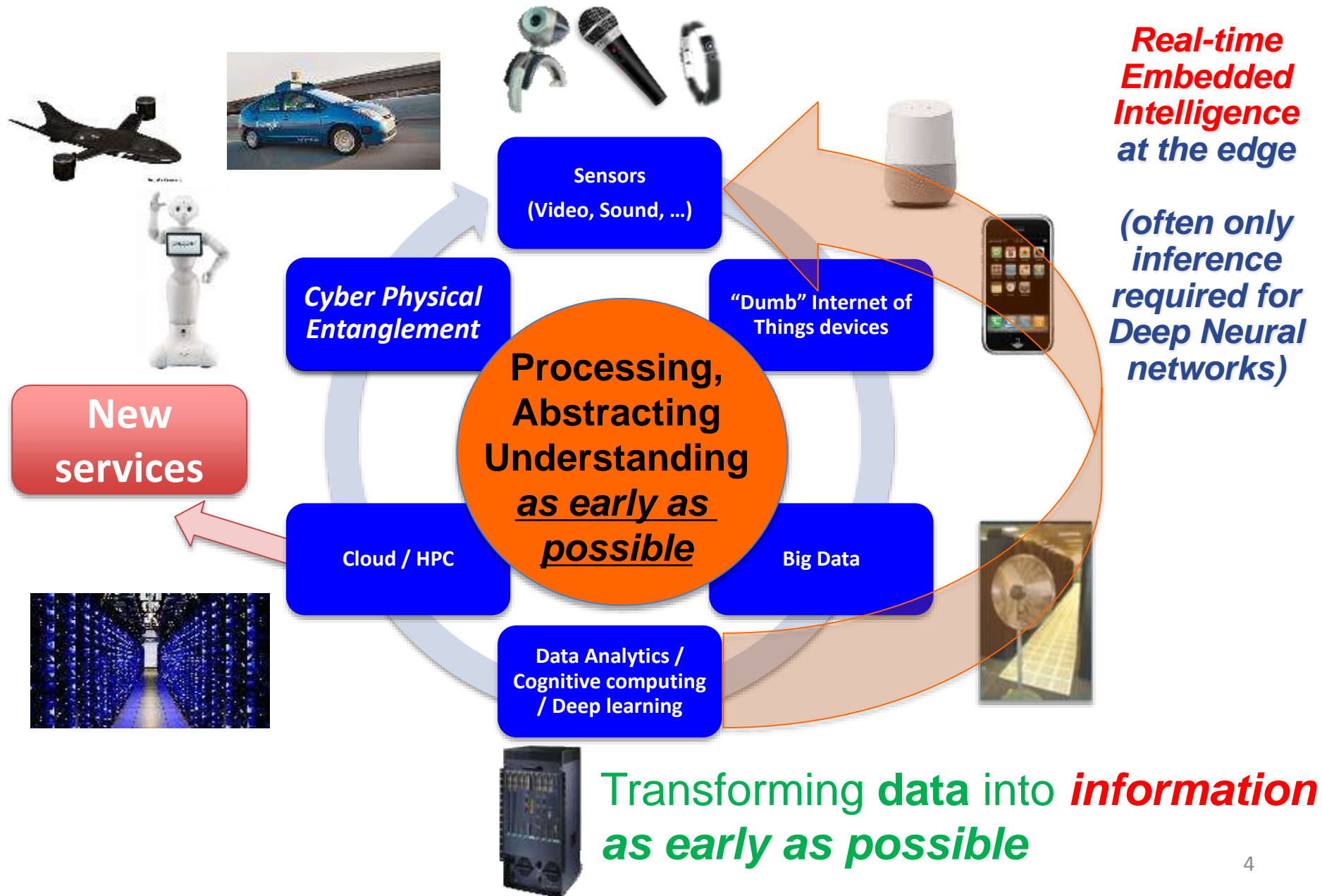
Editorial board:

Marc Durantou, Koen De Bosschere,
Christian Gamrat, Jonas Maebe,
Harm Munk, Olivier Zendra

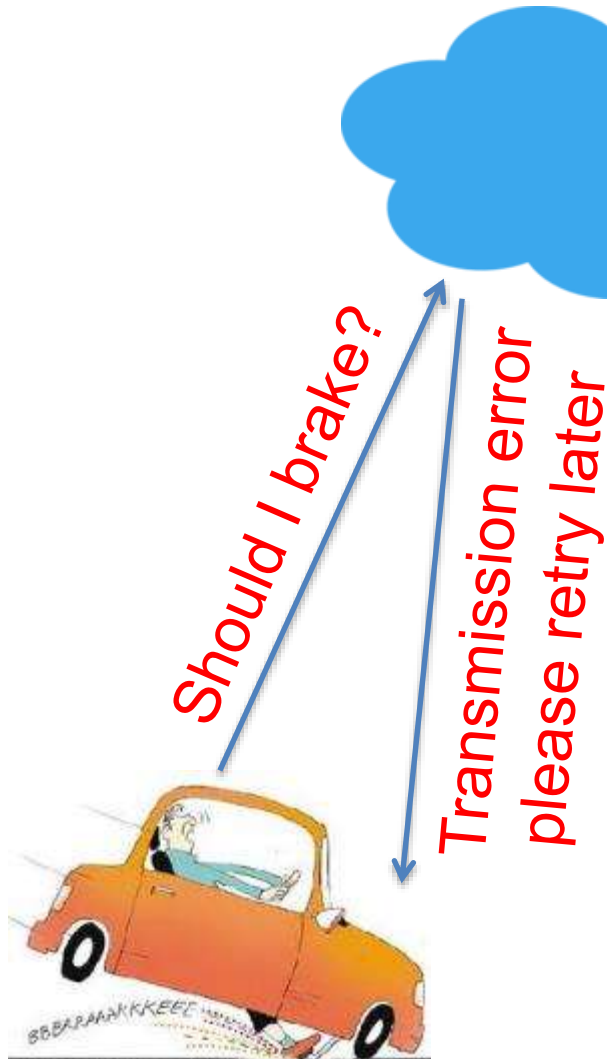
IOT SYSTEMS VERSUS CYBER-PHYSICAL SYSTEMS

There are many definitions of *Internet of Things* and *cyber-physical systems*, and a lot of controversy. We choose to define a *CPS* system as being characterized by having an actuator that directly impacts the physical world (a screen is not considered as an actuator in this definition), while an *IoT* system is distributed and composed of elements that communicate typically via the Internet. With our definition, CPS and IoT are not exclusive. For example, a self-driving car that is not connected and makes all its decision locally is a CPS device, but not an IoT device. It only becomes an IoT device (still being a CPS) if it is connected, e.g. to get maps from a server. A smart sensor transmitting the local temperature to a smartphone is an IoT device, but it is not part of a CPS. If it is connected to a thermostat that controls heating, the combination (i.e. the system composed of the sensor, the various servers and the thermostat) becomes a CPS (and the sensor is still a IoT device).

Computing Distribution for "Cognitive" systems



Embedded intelligence needs local high-end computing



System should be autonomous to make good decisions in all conditions

Safety will impose that basic autonomous functions should not rely on “always connected” or “always available”

Cloud and HPC cannot support many cyber-physical applications.

Embedded intelligence needs local high-end computing



Example: detecting elderly people falling in their home

Privacy will impose that some **processing** should be done **locally** and not be sent to the cloud.

Embedded intelligence needs local high-end computing



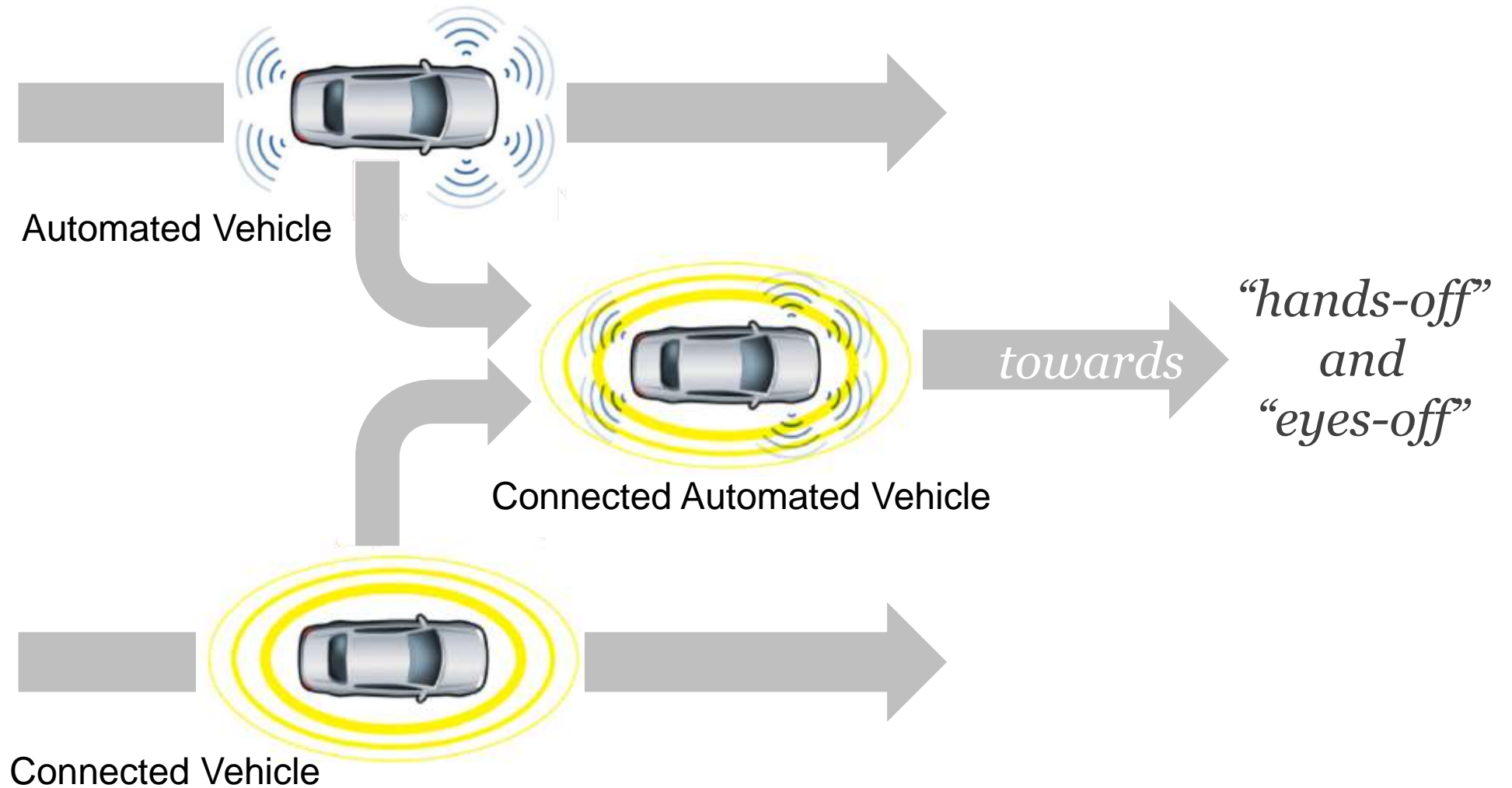
Dumb sensors



Smart sensors: image and
signal (voice) recognition

Bandwidth will require more **local processing** 7

CPS CHALLENGES IN TRANSPORT: SELF-DRIVING CAR



CPS CHALLENGES IN TRANSPORT: SELF-DRIVING CAR

ENVIRONMENT

SENSING - PERCEPTION - CONTROL/DECISION

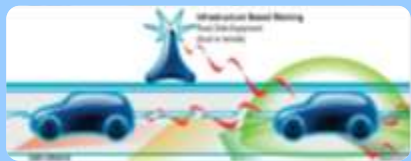
VEHICLE FEEDBACK

Source: Advanced Driver Assistance Systems (ADAS) by Bosch, Denso and Renesas, 2013

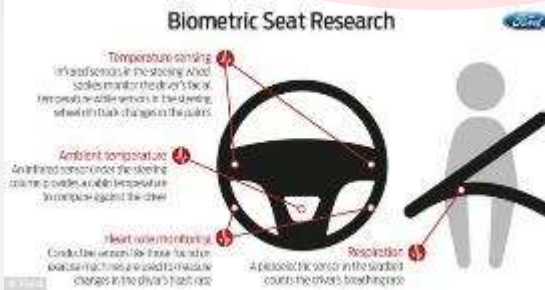
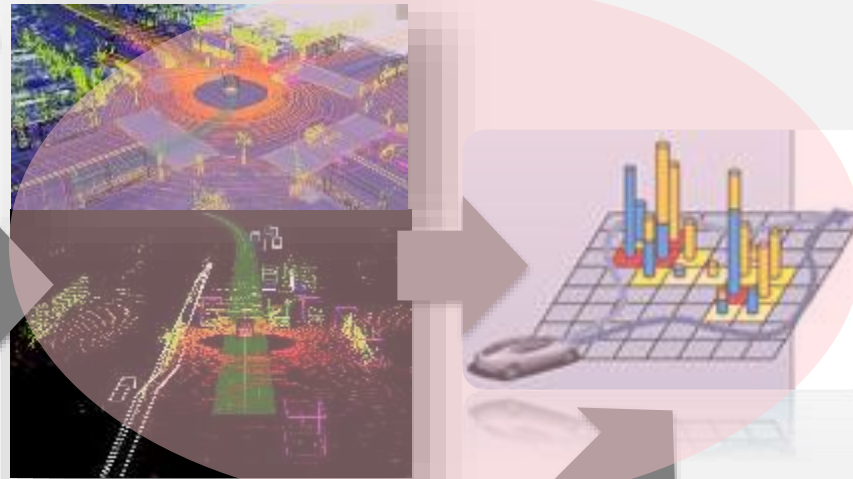


OBJECTS

V2X



Embedded AI



Source: Engineers develop a car that can monitor your HEALTH as you drive - and take over if you become ill or fall asleep, MailOnline, 2013



DRIVER INFO



X-by-WIRE

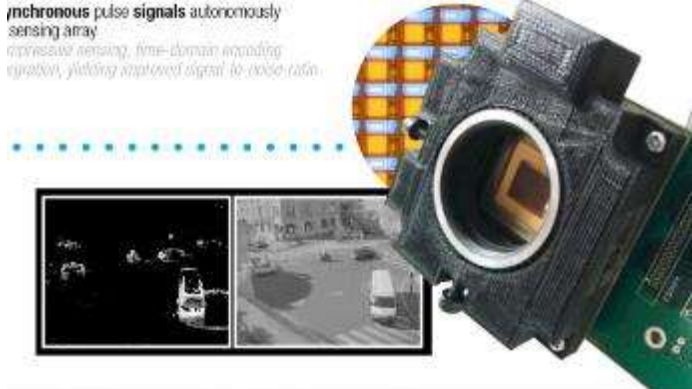
5 KEY ENABLING TECHNOLOGIES



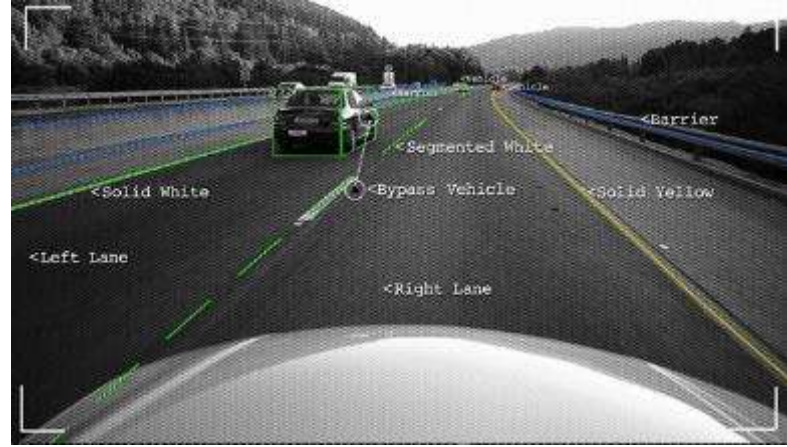
APPLICATIONS/ALGORITHMS

By continuously sweeping the sensor, asynchronous sensing algorithms process returning signals. These algorithms vary frequencies automatically depth and range output high-accuracy high-SNR exposures

Asynchronous pulse signals autonomously sensing array
compress sensing, time-domain expanding operation, yielding captured signal to noise ratio



Chronocam



Mobileye



Argo AI (Ford)



Cruise automation (GM)



Drive.ai



Nutonomy

And others....

COMMUNICATION CHALLENGES

Distributed application deployment

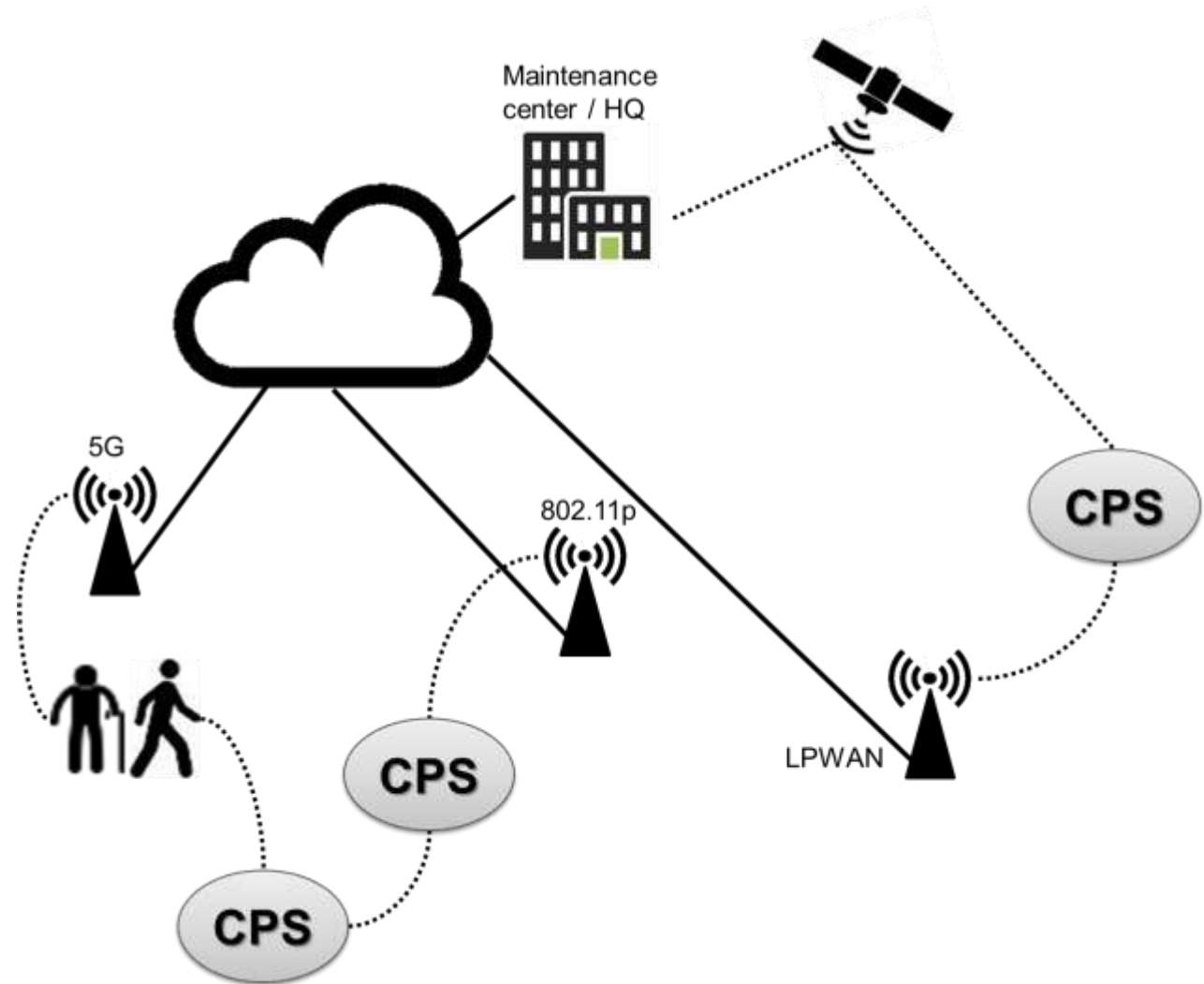
Dynamic network routing

Edge Computing in infrastructure

Multi-support wireless

Data privacy

IoT access authentication



► **CINGULATA: Secured data processing on unsecured servers**

"Homomorphic encryption (HE) allows to process encrypted information without decrypting it..."

- To design and translate the native application in HE domain
- To help integrate HE in the system architecture
- In-house & standard crypto-systems
- Application to industrial use cases:



Hidden navigation



Biometric



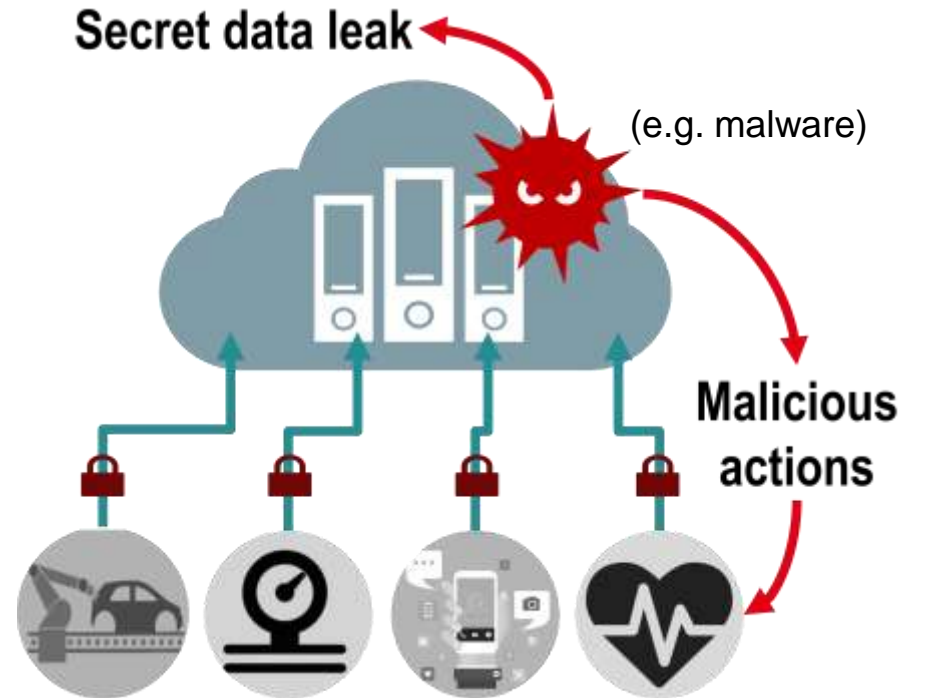
Intrusion detection



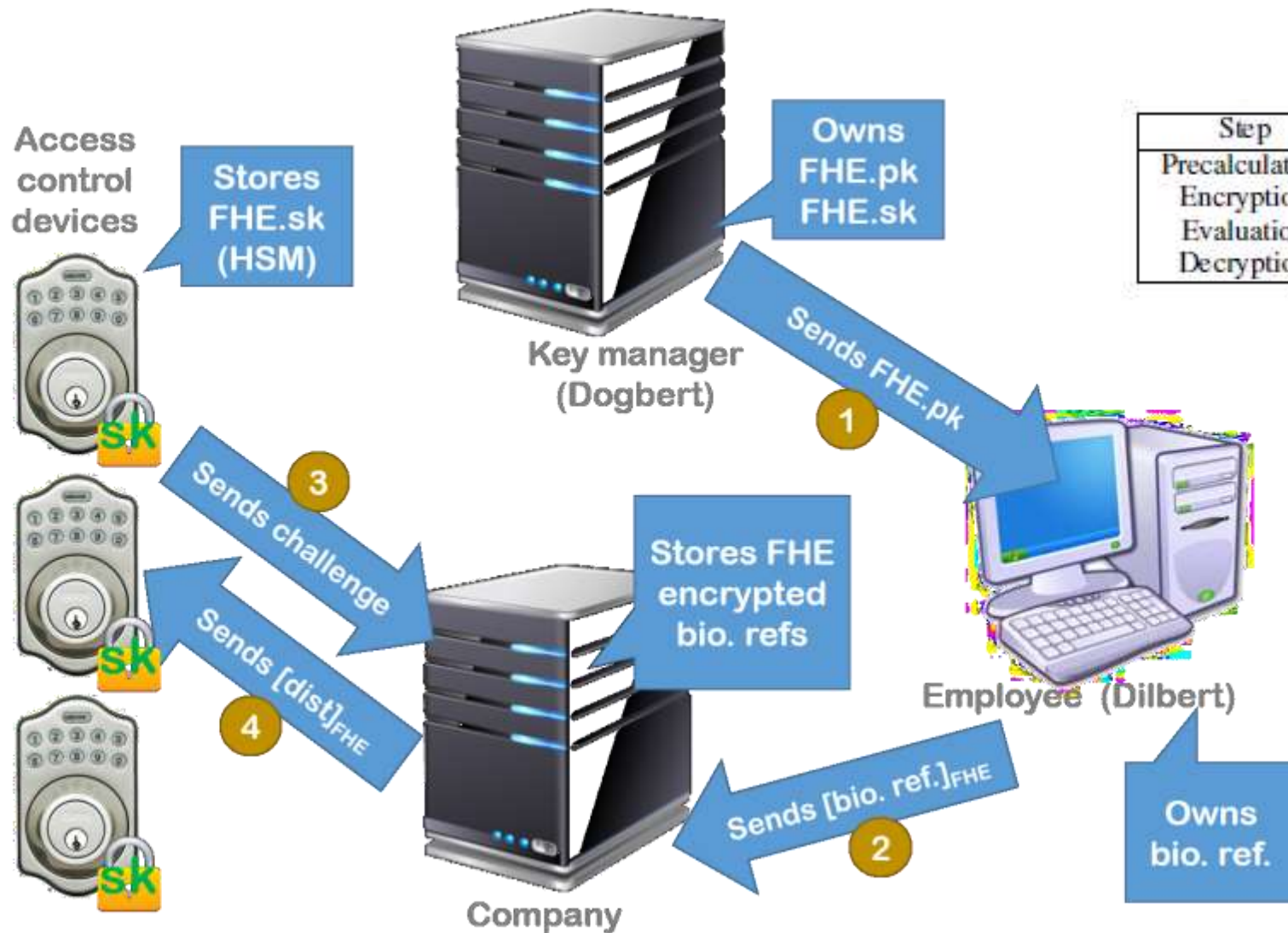
Healthcare



Smart Industry



ENABLING DATA PRIVACY : OVERVIEW



Step	When done	Where done	Duration
Precalculation	Offline	Dogbert	≈30 mins
Encryption	Enrollment	Dilbert	21 secs
Evaluation	Authentication	Company	<2 secs
Decryption	Authentication	Device	ε

TABLE I. PERFORMANCE SUMMARY

COMPUTING CHALLENGES

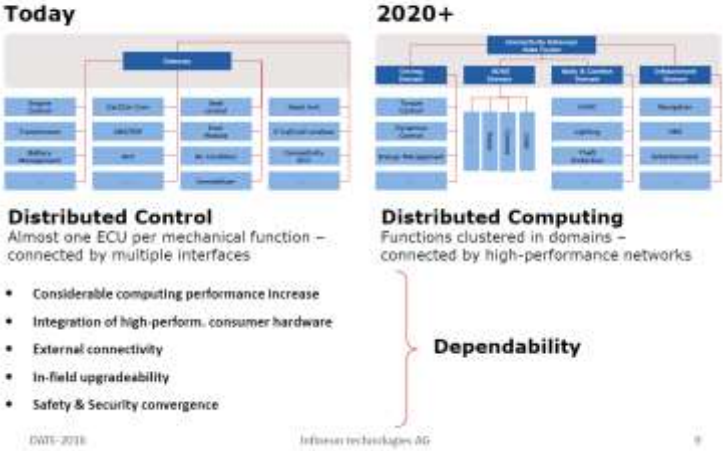
Decentralized vs. centralized

IT computing & communication technologies

SW & tools

Safety and Cybersecurity

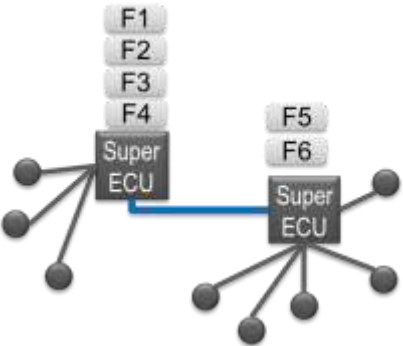
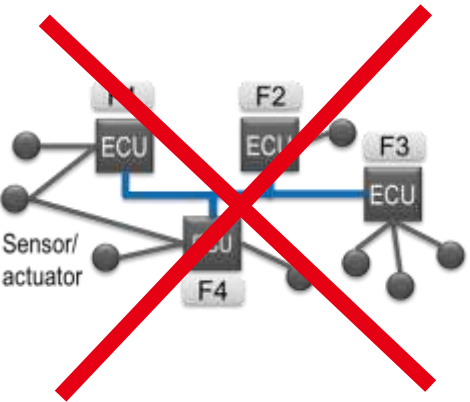
New requirements for the car electronics



Infineon, DATE2016



Siemens, RACE project



Tesla, Model S



Audi, zFas

COMPUTING CHALLENGES

Decentralized vs. centralized

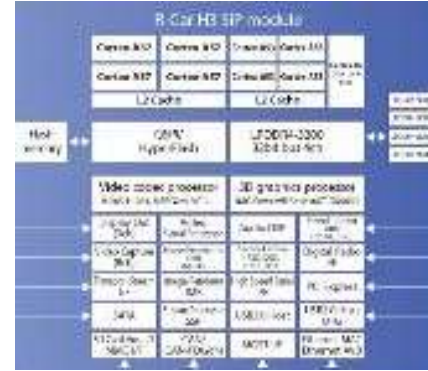
IT computing & communication technologies

SW & tools

Safety and Cybersecurity



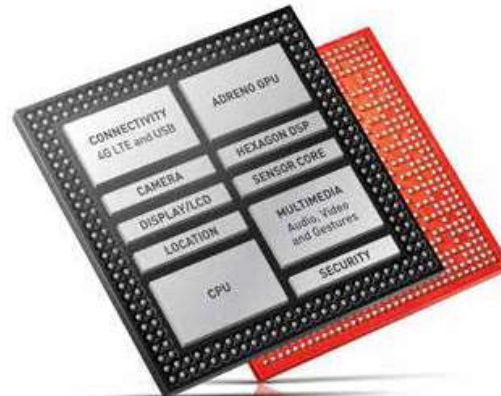
**MOBILEYE
(EyeQ5)**



**RENESAS
(R-Car V3H)**



**NVIDIA
(Drive PX2)**



**QUALCOMM
(Snapdragon 820)**

**INTEL
(Xeon D1529)**



**KALRAY
(MPPA Bostan)**

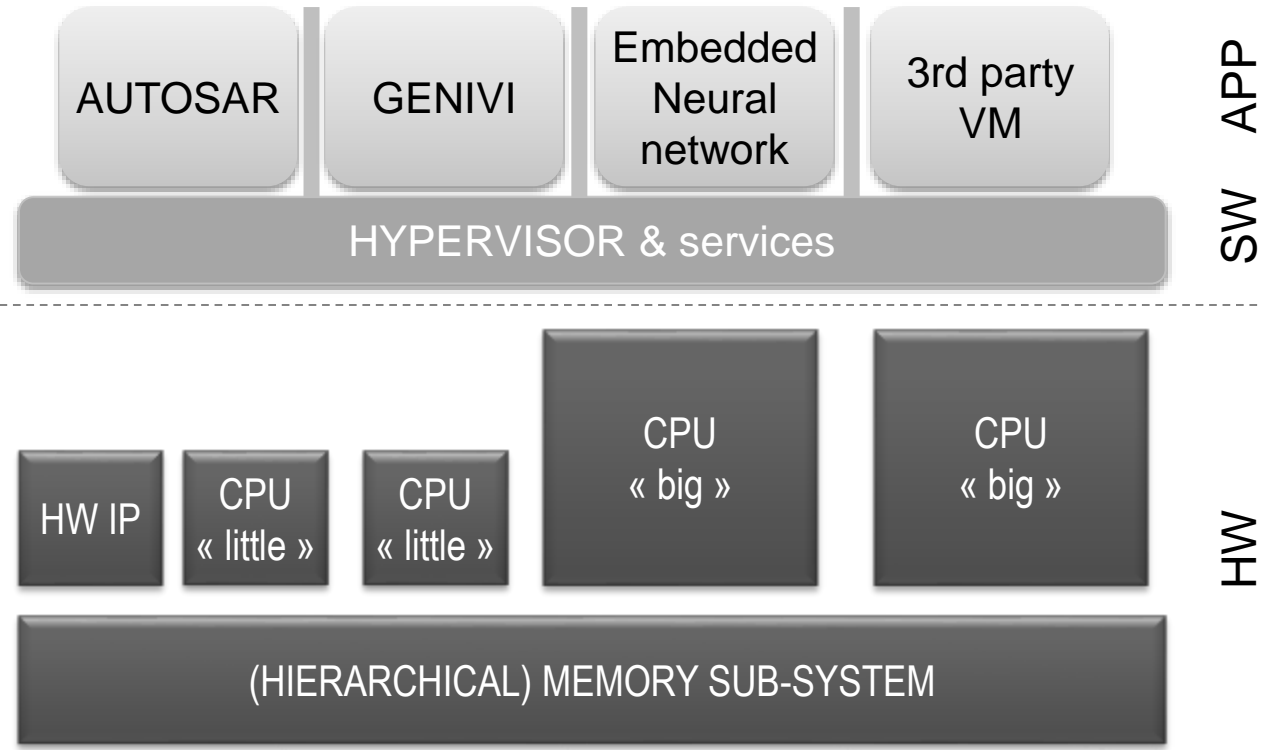
And others...

COMPUTING CHALLENGES

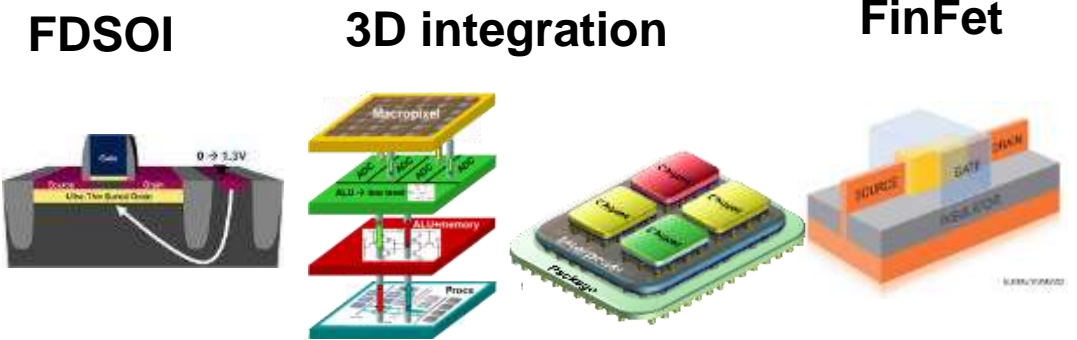
- Decentralized vs. centralized
- IT computing & communication technologies
- SW & tools
- Safety and Cybersecurity

Tools

Safety & cybersecurity



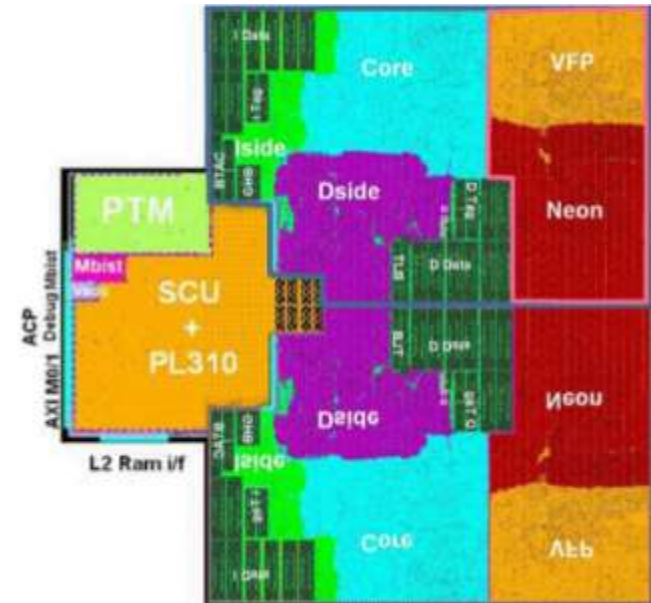
APP
SW
HW



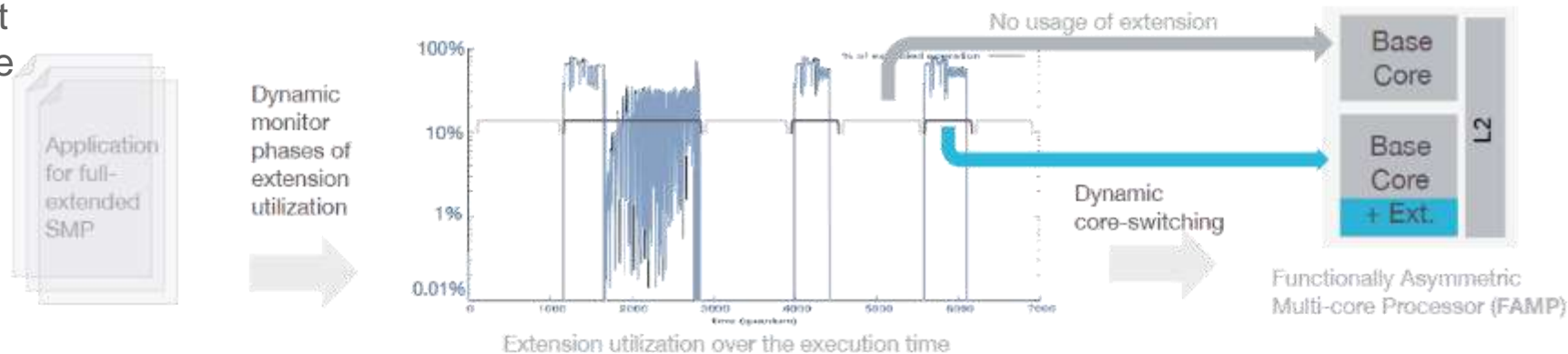
TECHNO

DYNAMIC MANAGEMENT OF FAMP

- In SMP architectures, extensions accelerate some specific tasks from 4x to 1000x
 - But are used less than 5% of time
 - May consume up to 25% of the processor area
- Functionally Asymmetric Multi-core Processor (FAMP)**
 - Objectives are
 - To maintain a reduced silicon area
 - To limit performance degradation
 - To reduce the energy consumption
 - Techniques
 - Limits the use of costly extensions for critical sections
 - Optimizes task placement according to performance



CORTEX A9 dual core Floorplan
From Osprey – 1.9W TDP 2GHz (6.7mm²)

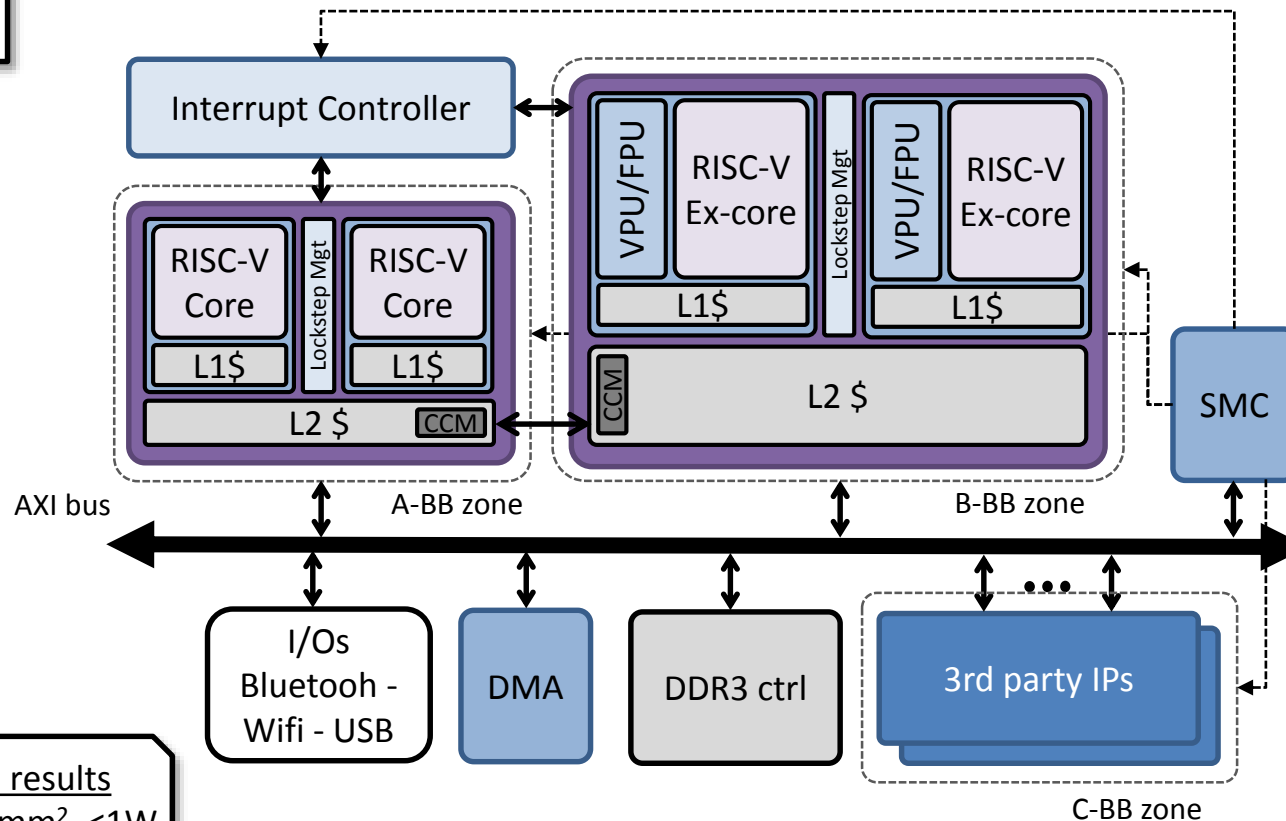
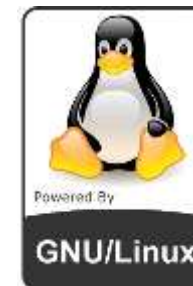


HETEROGENEOUS MULTICORE DESIGN

32/64-bit RISC-V cores
Single-issue in-order
2008 IEEE-754 FPU
SESAM simulator
L2 cooperative caching
Overlapping ISA



SMP Linux
GCC/GDB



28nm FDSOI results
800MHz, 2mm^2, <math><1\text{W}</math>

Smart Management
Controller (SMC)
*Ageing, body-bias,
temperature...*

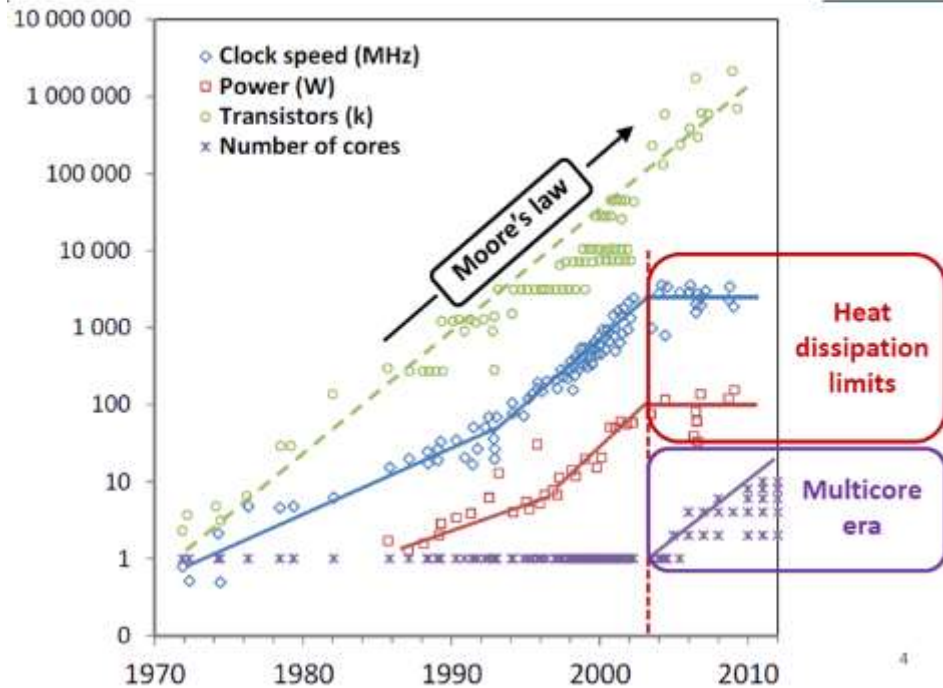




PART II: FDSOI & RELIABILITY (OVERVIEW)

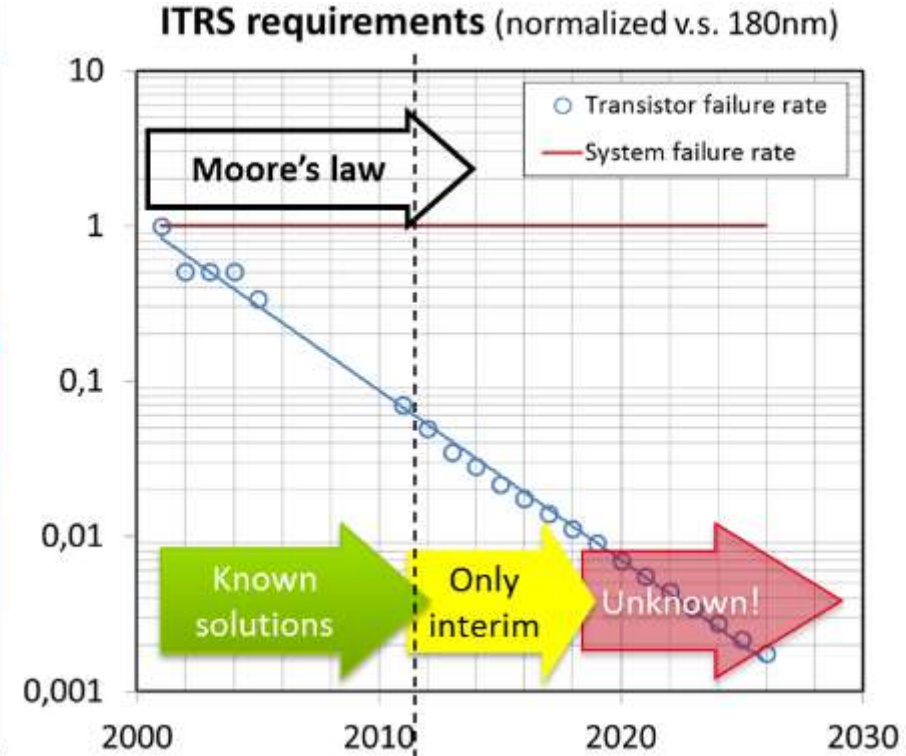
CONTEXT

Source: V. Huard, IRPS2013



Multi-/many-core

Source: ITRS PIDS2000&2011



Ageing/wear-out

Bulk technology limits due to short channel effects

Fully-Depleted SOI transistor

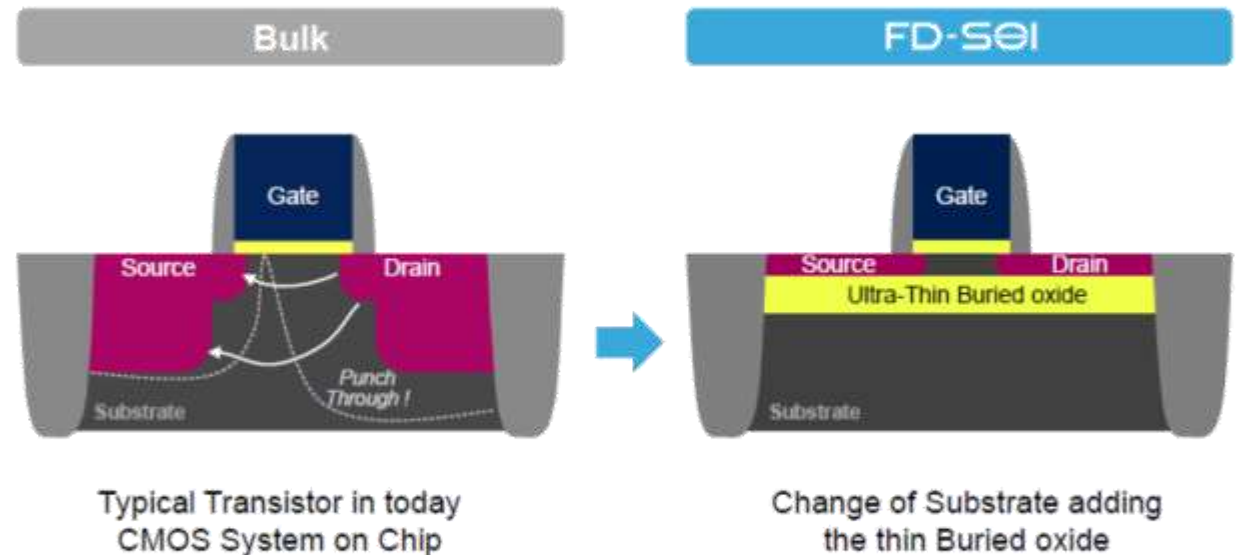
Adjustable transistors V_t through body biasing (BB)

Improved variability

Low cost process technology

Fully compliant with IP designs already available in Bulk

FD-SOI Transistor 2



Improving power Efficiency – Bringing high flexibility in SoC integration



Source: "Advances in Applications and Ecosystem for FD-SOI technology", Philippe Magarshack, STMicroelectronics

KEY AUTOMOTIVE CHALLENGES

Constraints

Temperature

*Large spatial & temporal gradients in a vehicle
Heat dissipation solutions are limited
Must-be compliant to AEC Q100 standard*

Reliability

*Autonomous vehicle reinforce the mandatory
requirement of continuity of service
10 to 1 FIT objective (are still?) over 30 years*

*Low Power
& Energy*

*1. Thermal dissipation ($\sim 10^1 W$)
2. When vehicle is switched-off, some electronics
remain powered-on (Door opening, OTA,...)*

Performance

*ADAS & autonomous vehicle (IA) → growth demand
for embedded computing solutions in vehicle*

Needs

FDSOI enables 32% speed boost compared to Bulk at 1V (no BB)

Adaptive body biasing: wider bias range available v.s. bulk

BB can be adjusted dynamically by application

Vt reduces → +10% performance (FBB) v.s. no BB

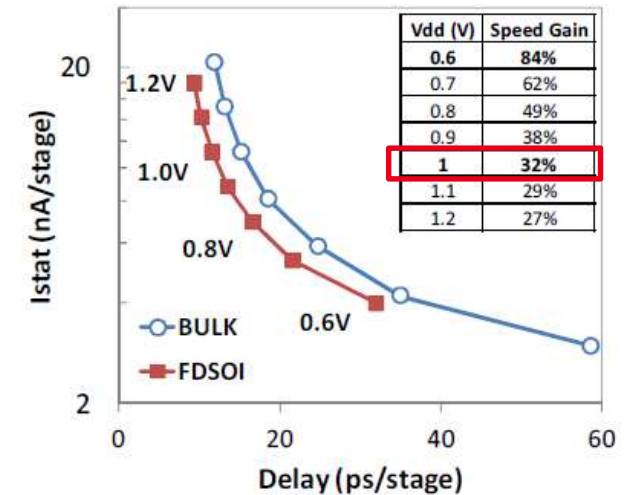
Vt increases → static power /2 (RBB) v.s. no BB

*FBB → High performance application
RBB → Low power application*

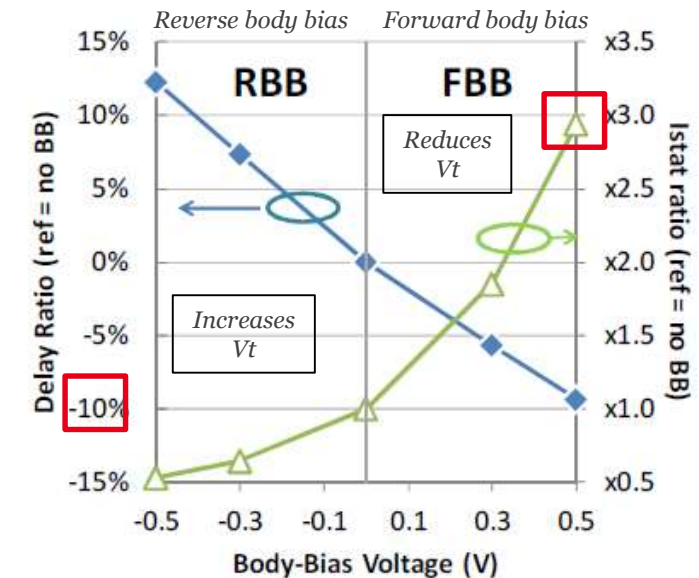
Performance-power tradeoff => one technology covering two application objectives

... But increases by 3 static power

... But decreases performance by 10%



Measurements on a ring oscillator



Bulk technology suffers from ageing and radiation

Ageing causes variability during operation



Circuit performance decreases

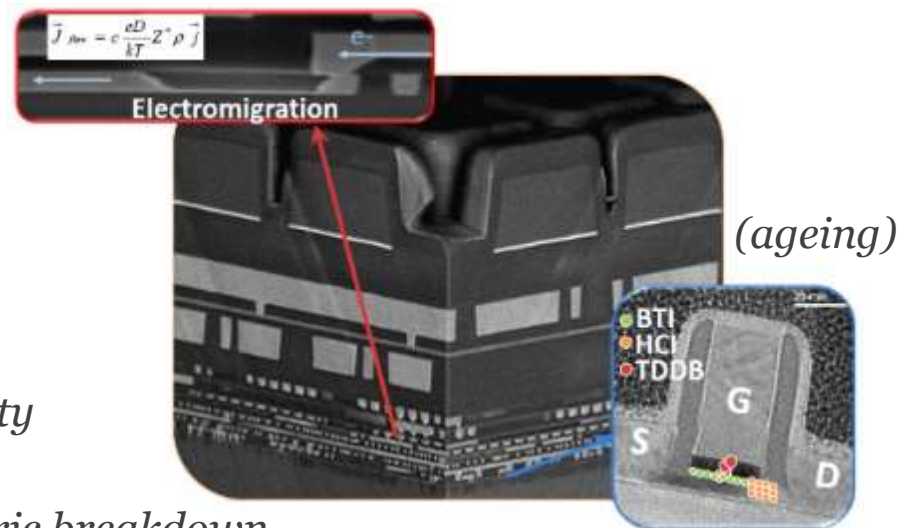
Radiation causes bit flipping in storage cells



SRAM reliability decreases

FDSOI does not cause new failure mechanisms

FDSOI improves transistor reliability v.s. Bulk



BTI: Bias temperature Instability

HCI: Hot Carrier Injection

TDDB: Time-dependant dielectric breakdown

*Better neutrons immunity v.s. Bulk
upset rate < 10FIT/Mb*

Single Event Latchup immunity

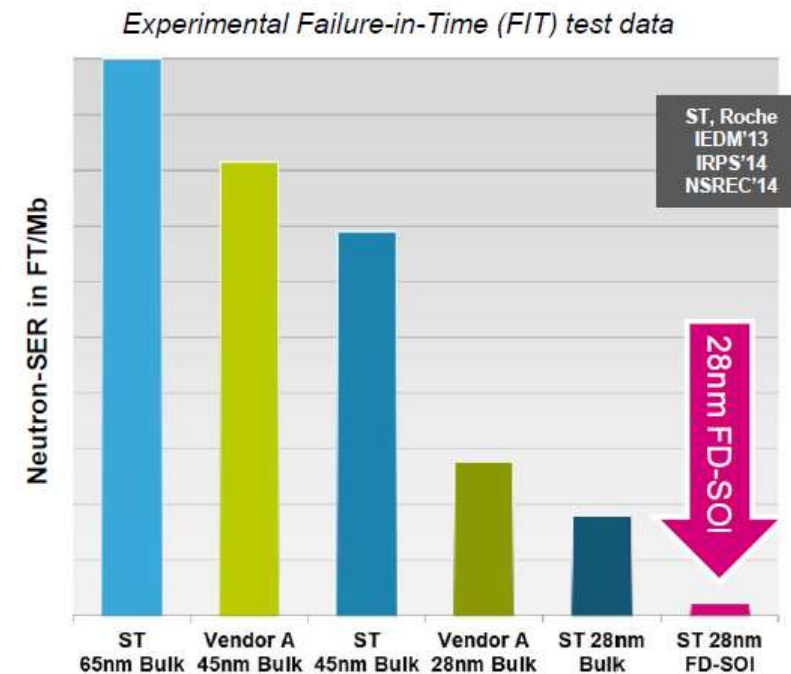
*Alpha particle quasi-immunity.
upset rate < 1 FIT/Mb*

*No need for ultra-pure
Alpha packaging*

*Very small error clusters
(99% single bits)*

*Single error correction
codes are sufficient*

*FDSOI: 100× to 1000× more reliable than Bulk
at Sea-level and Space (also better than FinFET)*



Gain w.r.t. BULK	UTBB FD-SOI	FinFET
Alpha	1000×	15×
Neutron	100×	10×
Latchup	immune	not reported
Multiple Cell Upsets	99% single bit max. 2 cells	max. 4 cells worse according to INTEL

ST, Roche, CISCO SER workshop, Oct'14

TSMC, Farg, CISCO SER workshop, Oct'14

Source: Ph. Roche, ST Microelectronics

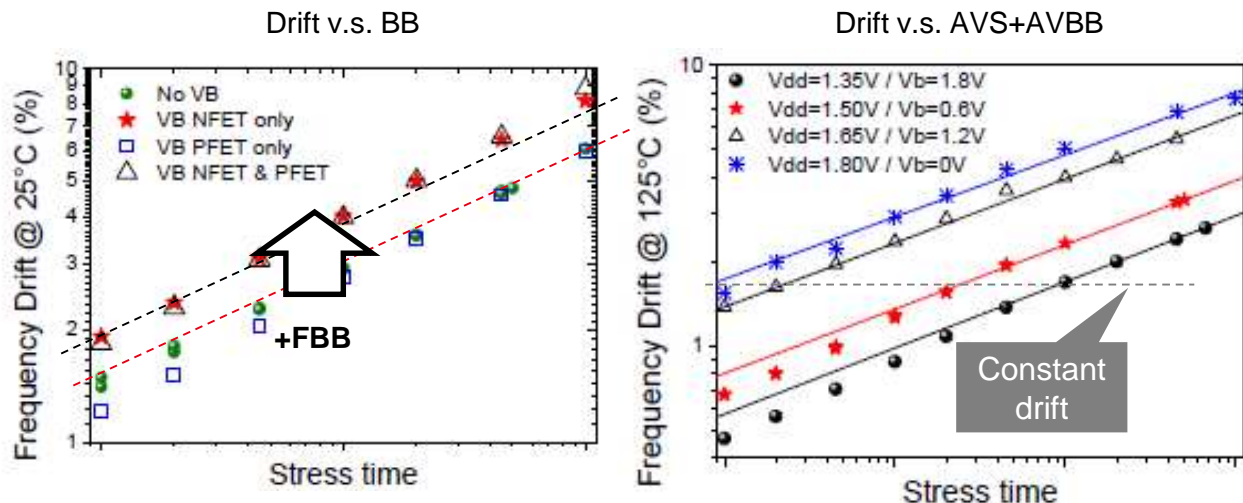
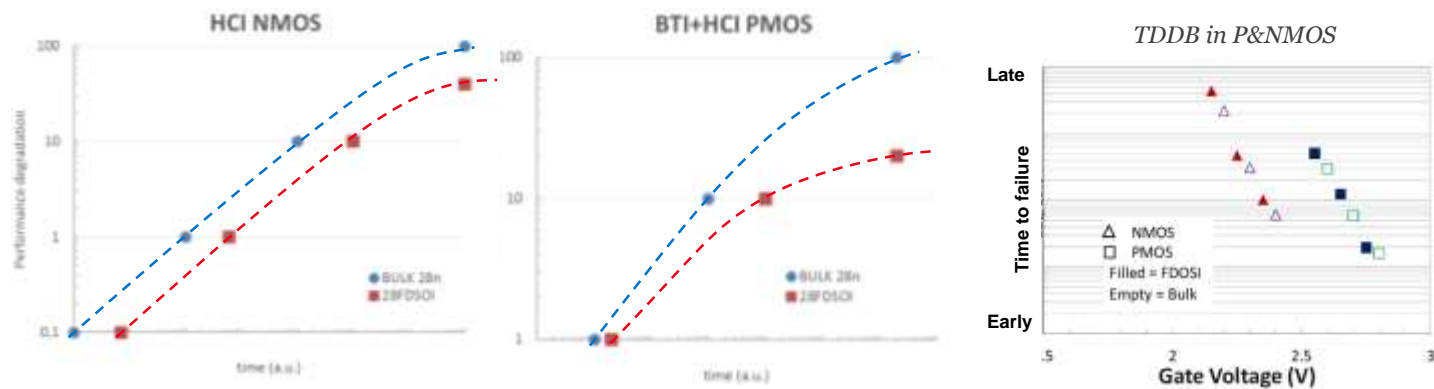
HCI + BTI: degradation is 2× to 7× slower than in Bulk

TDDDB: degradation is 2× slower than in Bulk

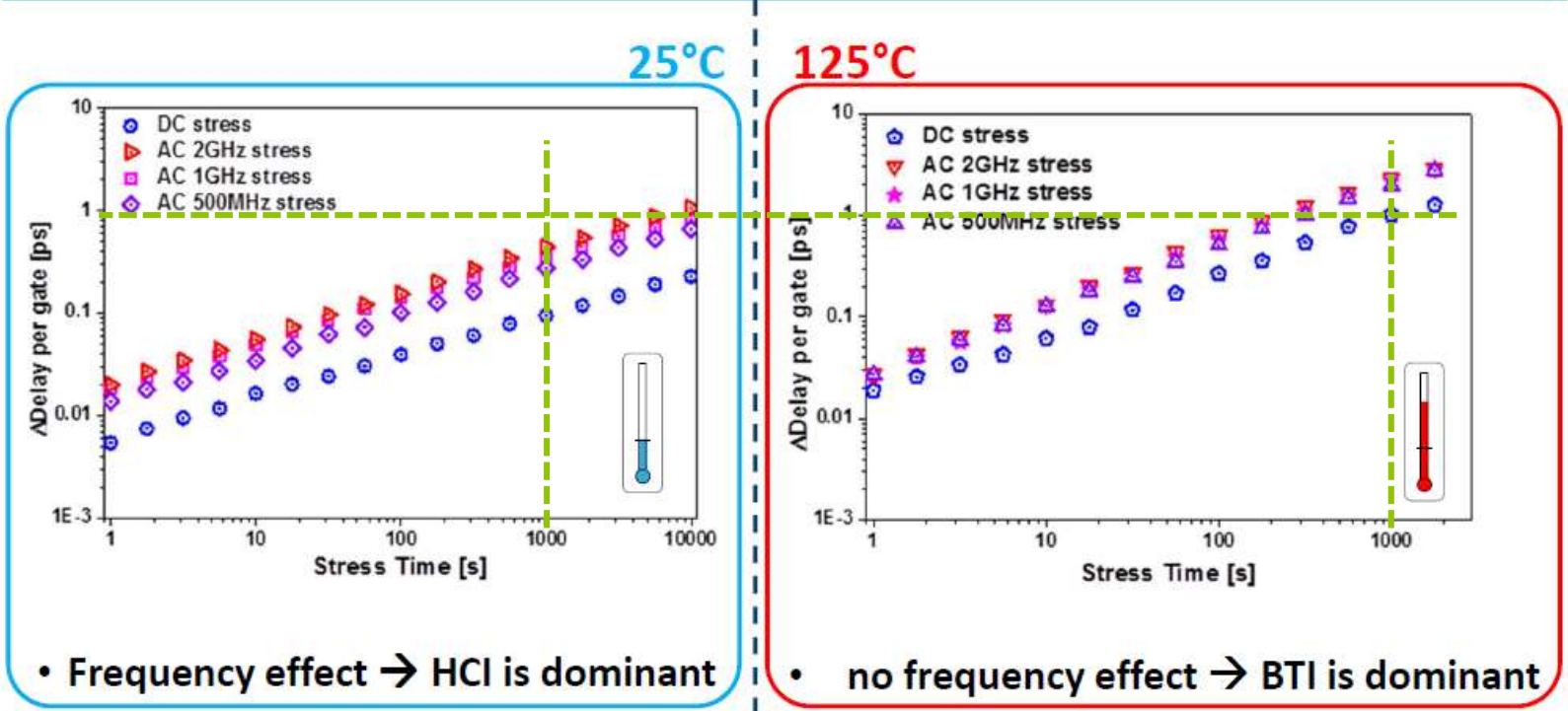
@VDD, FBB increases degradation v.s. no BB

AVS+ABB can fully compensate ageing with performance increase

Although FBB is active, ageing also depends on Workload significantly



Circuit Reliability Impact on circuits



Worst-case degradation @ 125C dominated by BTI

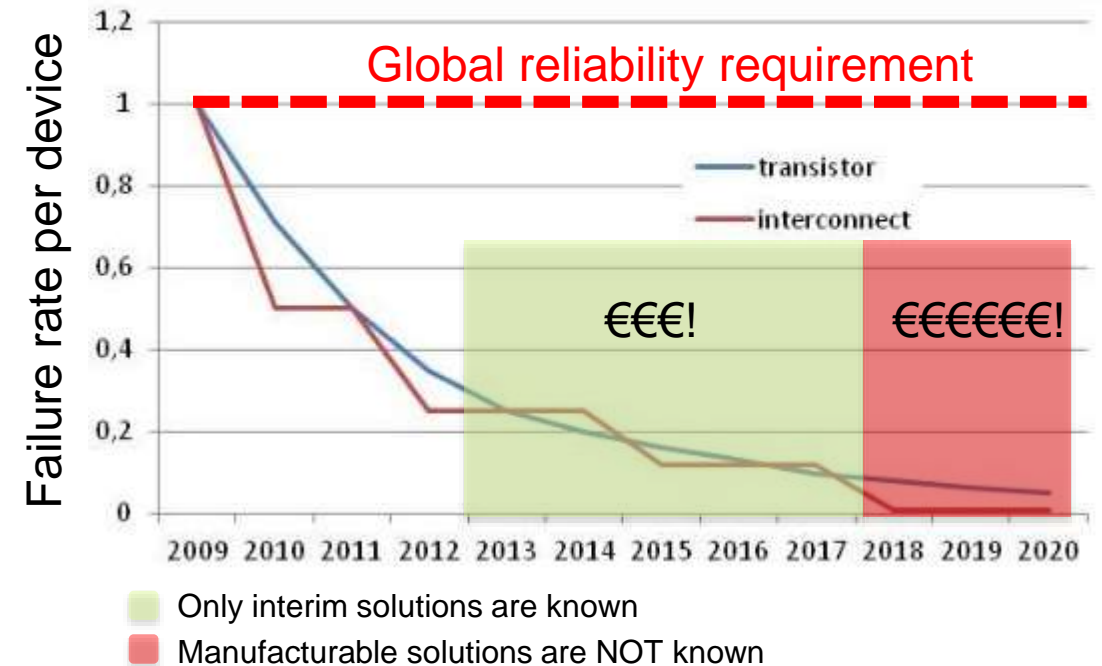
Source : Saliva et al., DATE, 2015

CONCLUSION

- Embedded applications need >100 GOPS/s
- Multi-Processor (many-core) System on Chip
- Transistor shrinking
 - (+) integration
 - (+) performance
 - (+) power
 - (-) variability
 - (-) leakage
 - (-) temperature
 - (-) reliability

- Non-uniform workload distribution → non-uniform distribution of failure rate
- Reliability sign-off is already done in back-end
- Industry needs solutions at higher abstraction levels

[ITRS, PIDS figures, 2009]





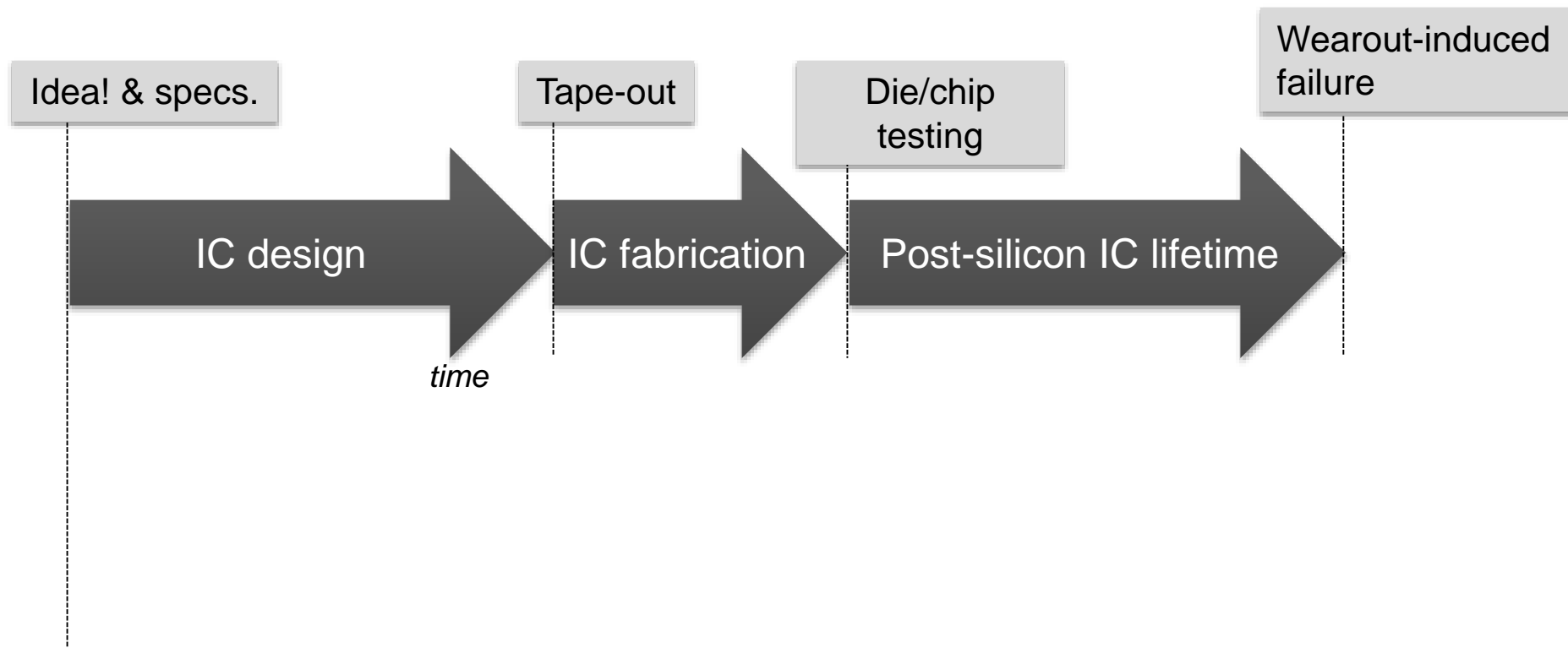
PART III: AGEING AT ARCHITECTURE LEVEL

Ageing device models?

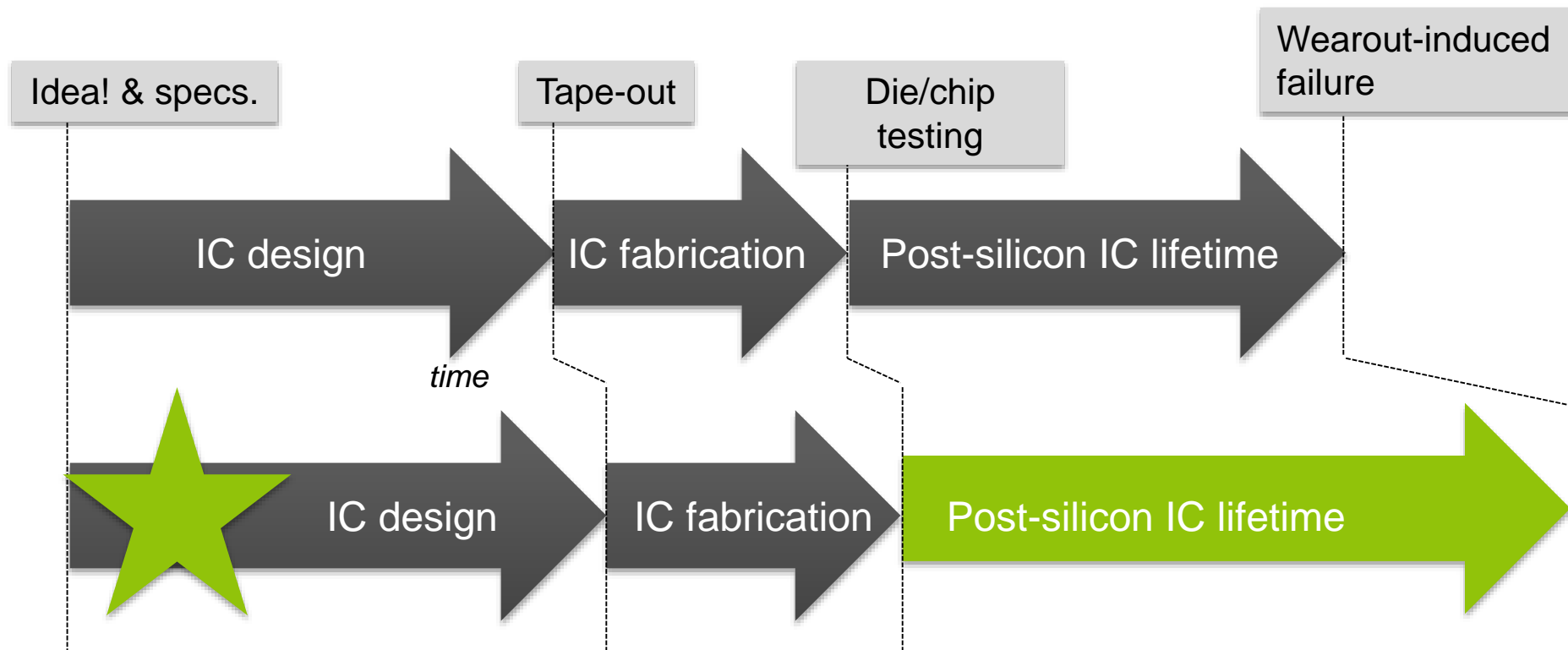
How scaling from device to gate, then circuit?

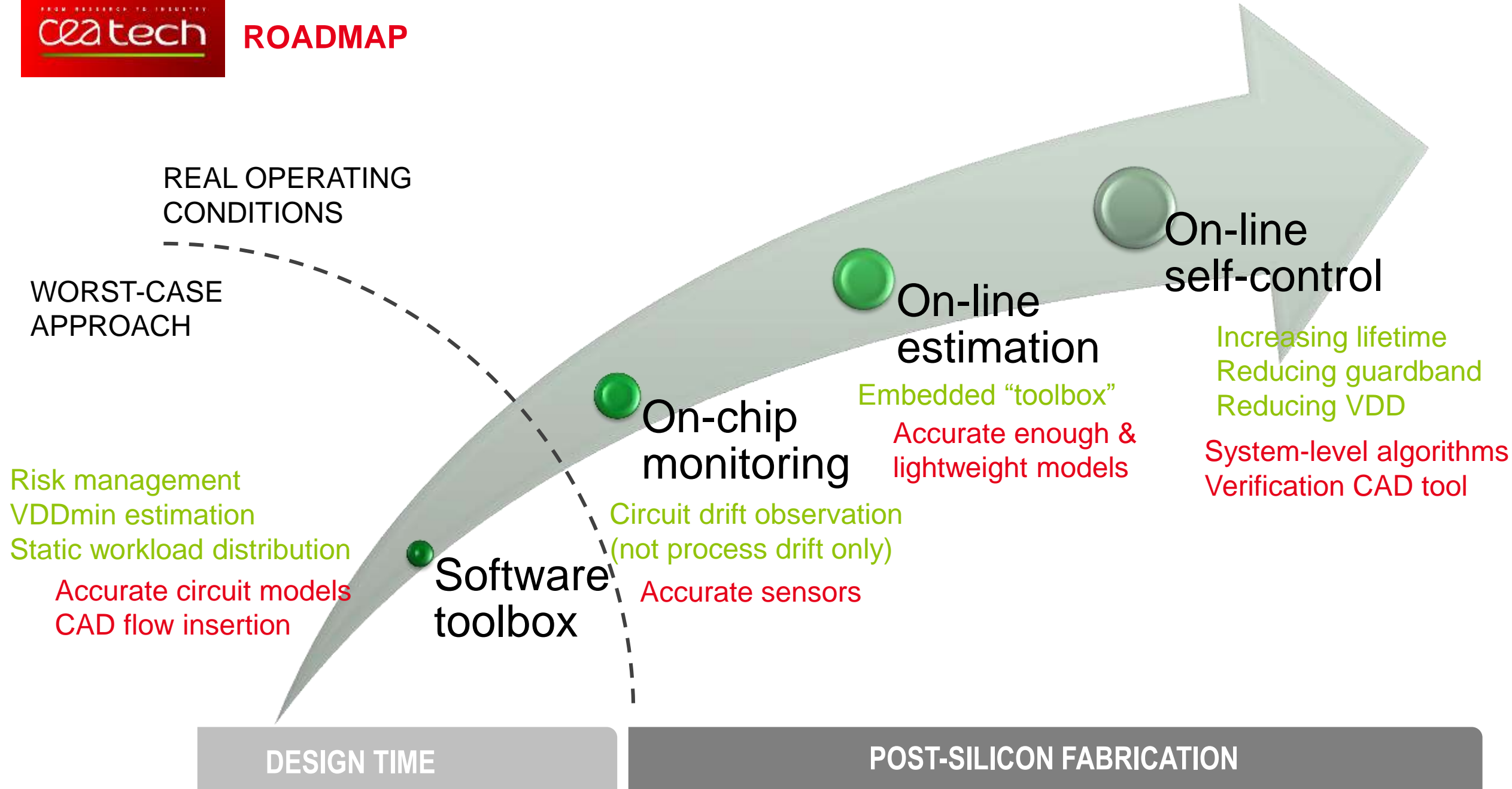
**How to simulate at Register-transfer?
at Transaction level?**

To understand and prevent ageing induced defects in digital ICs as early as possible at design time



To understand and prevent ageing induced defects in digital ICs as early as possible at design time





Process-level (test & characterization)

DfR Solutions
reliability designed, reliability delivered



etc...

TCAD & Device-level

DfR Solutions
reliability designed, reliability delivered



etc...

Gate-level



etc...

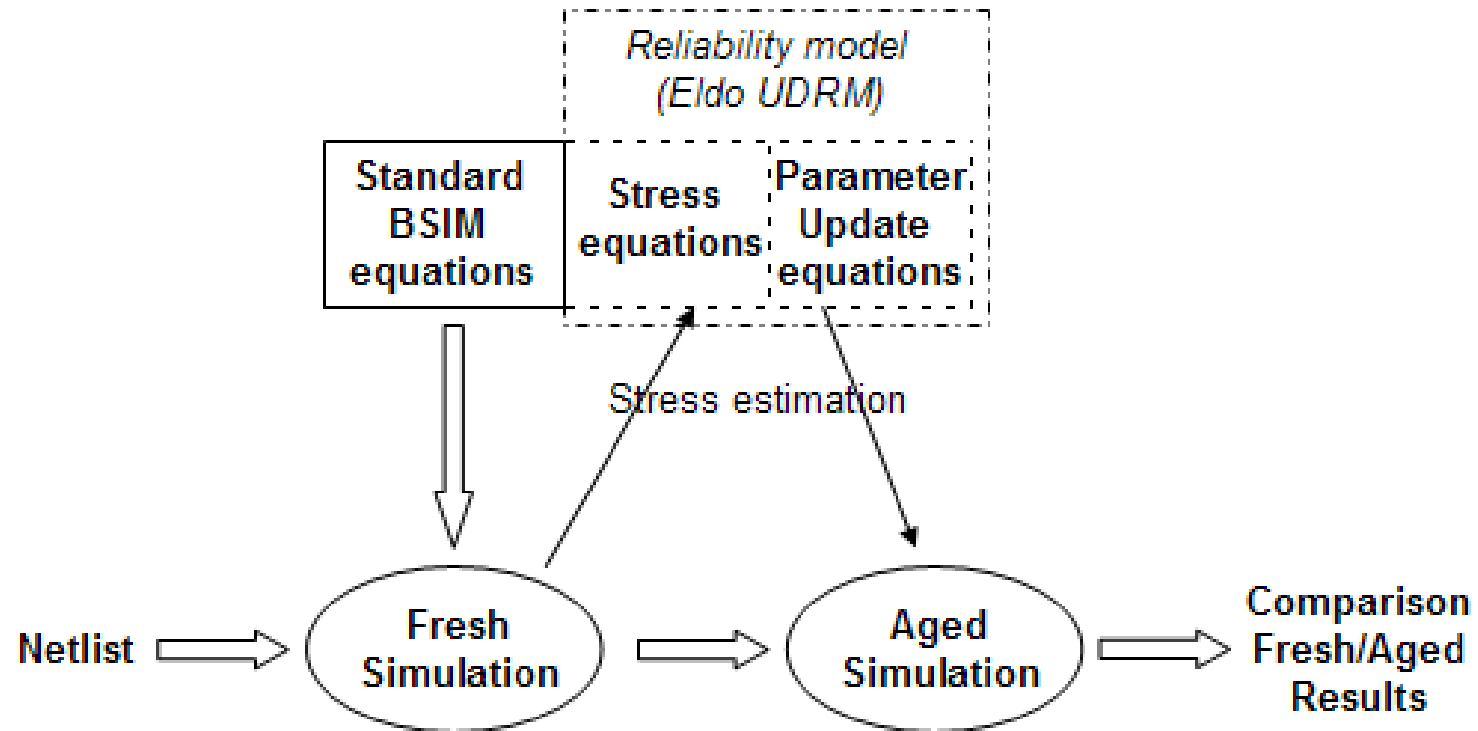
RELIABILITY SIMULATION: OVERVIEW

Level	Reliability monitors	Main factors	Existing solutions (not exhaustive...)	+/-
Device	Shift of electrical parameters	<ul style="list-style-type: none"> •Transistor/wire geometry •Currents, voltages, T°C, simulation time 	<ul style="list-style-type: none"> •BERT (Berkeley) •HOTRON (TI) •RelXpert (Cadence) •ELDO Premier (MentorGr.+STMicr.) •Etc. 	<ul style="list-style-type: none"> + Accurate (sign-off) + Silicon proof - Slow simulation - Too late for design exploration
Gate	Shift of propagation delay	<ul style="list-style-type: none"> •Cell nature •Stress time (toggling rate or duty cycle) •T°C 	<ul style="list-style-type: none"> •ILLIADS (Univ. of Illinois) •GLACIER (BTA techn.) •In-house solutions (STMicroelectronics, Infineon) •OFFIS 	<ul style="list-style-type: none"> + Silicon proof + Early reliability projections - Too slow for design exploration
Architecture	MTTF, CFR Delay degradation	<ul style="list-style-type: none"> •Area •Workload •Power •T°C •Process params 	<ul style="list-style-type: none"> •RAMP (IBM+Univ. of Illinois) •Univ of California San Diego+EPFL •RAAPS, DAPHNE (CEA LIST) •DAC&DATE2016 	<ul style="list-style-type: none"> + Generic models + Fast design exploration (virtual prototyping) - Model accuracy?

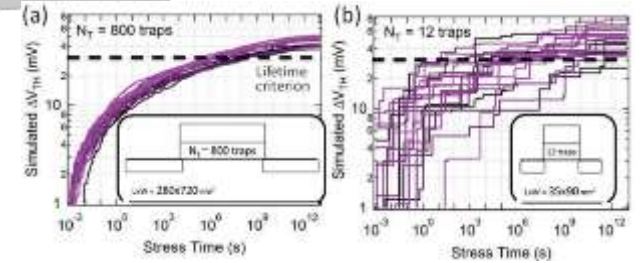
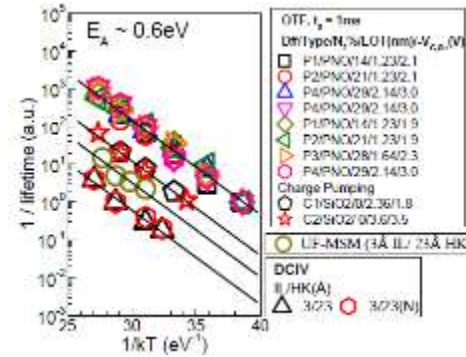
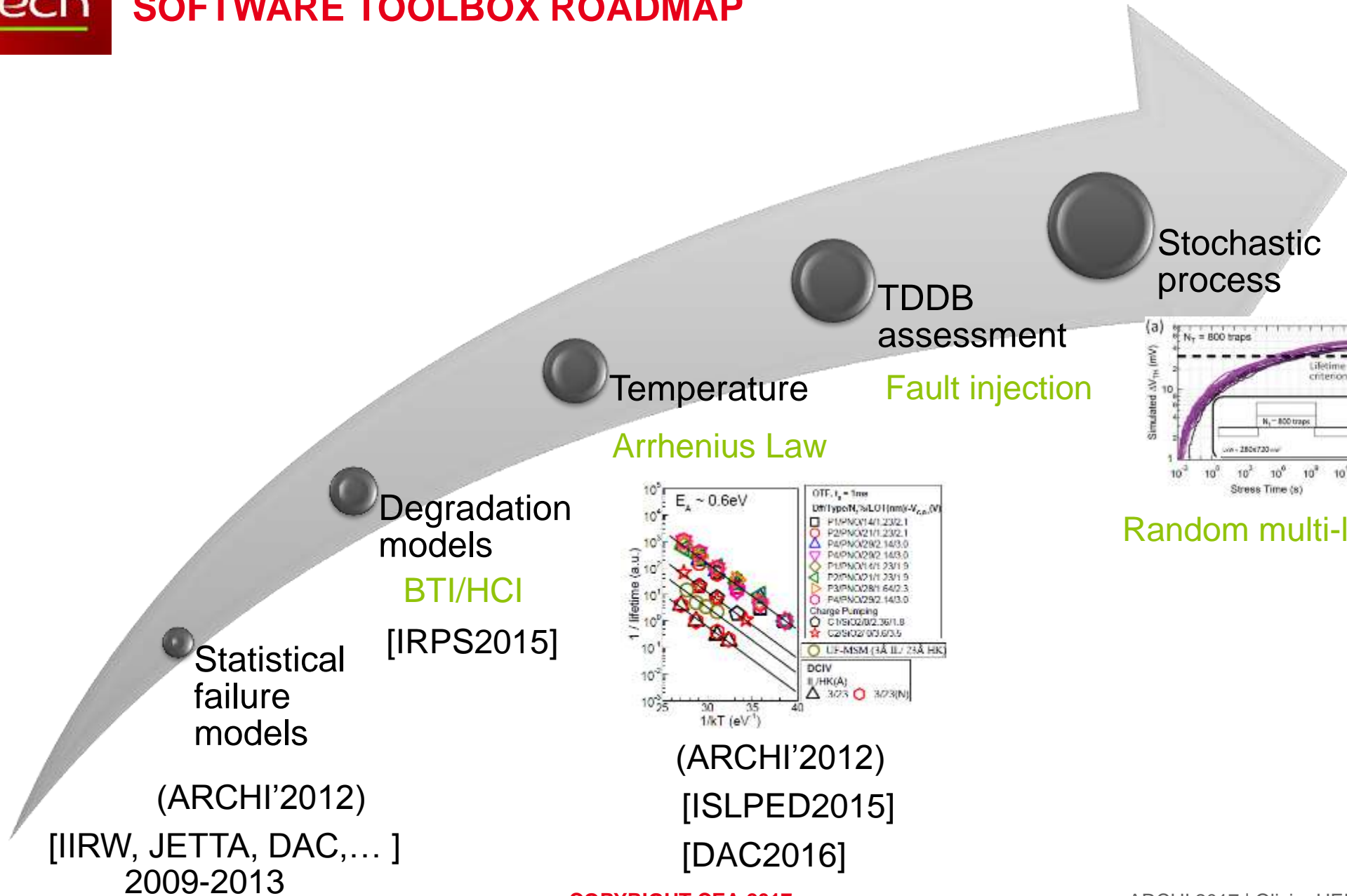
MTTF: Mean Time To Failure
CFR: Cumulative Failure Rate

RELIABILITY SIMULATION

- E.g. Eldo Premier flow for reliability simulation



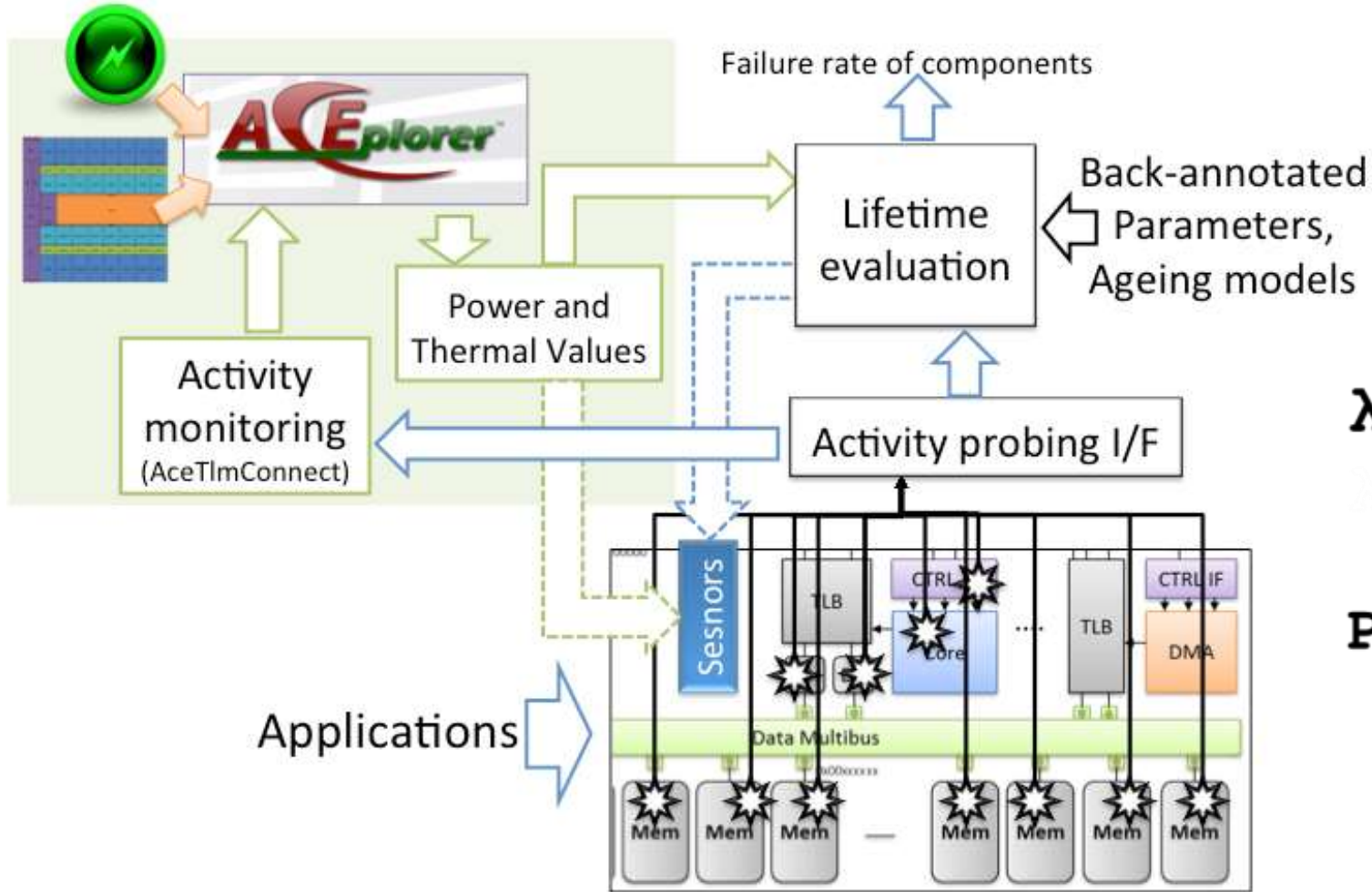
SOFTWARE TOOLBOX ROADMAP



FAILURE RATE ESTIMATION AT TRANSACTION LEVEL IN A NUTSHELL

CATRENE/RELY
[DAC2014]

POWER-TEMPERATURE SIMULATION (Intel/Docea Power)



Applications

SESAM Simulator

$$\lambda_{comp} = \sum (\lambda_{dev}(i))$$

$$= N_{dev} * Pr_{stress} * \lambda_{dev}$$

$$Pr_{stress} \sim (1 - s * P_{dyn}) / 2$$

S: fitting param

N: transistor count

λ_{dev} : device failure rate [JEDEC]

RELIABILITY SIMULATION AT REGISTER TRANSFER LEVEL

- Path propagation delay

T_p : path propagation delay

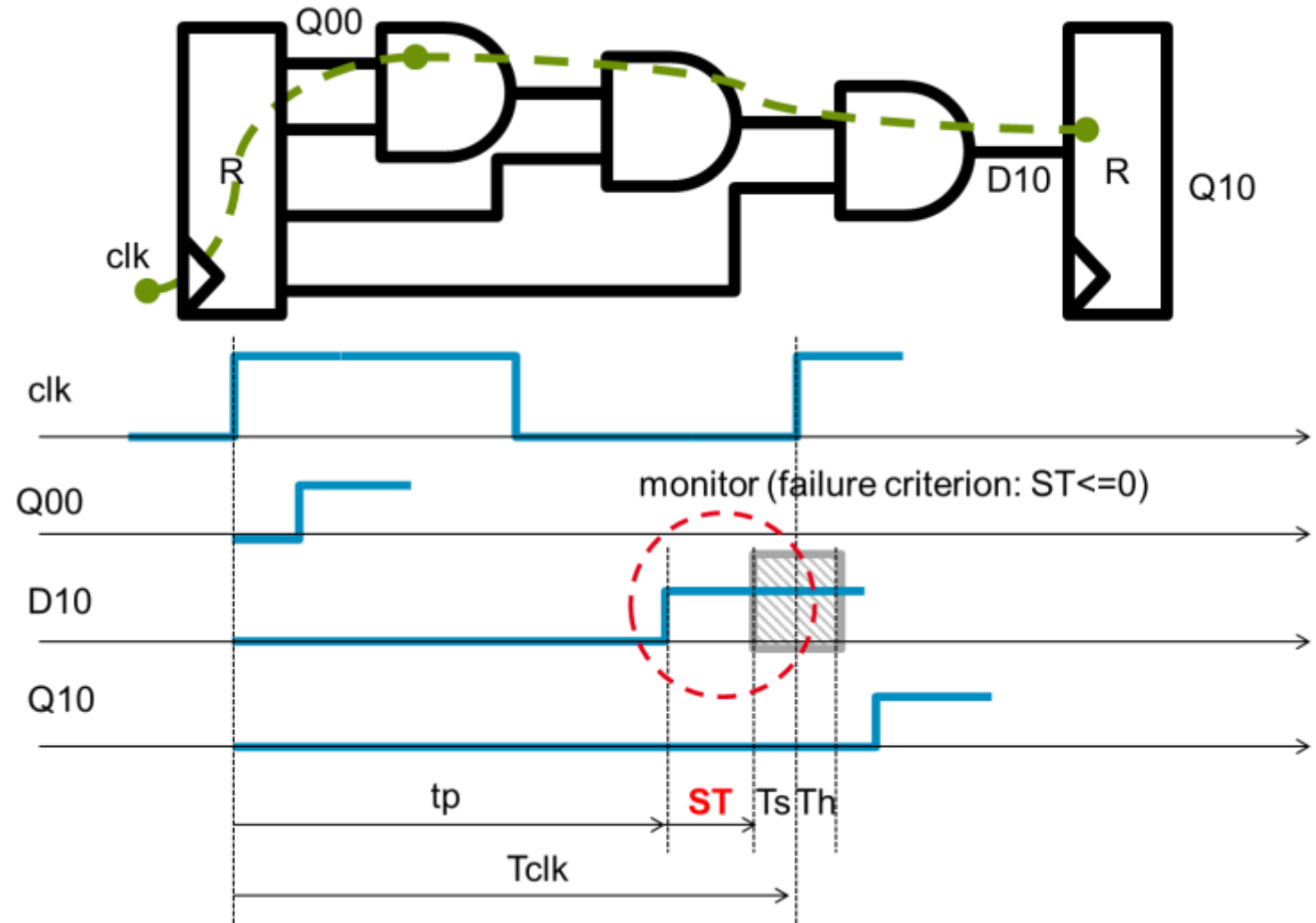
T_s : setup time

T_h : hold time

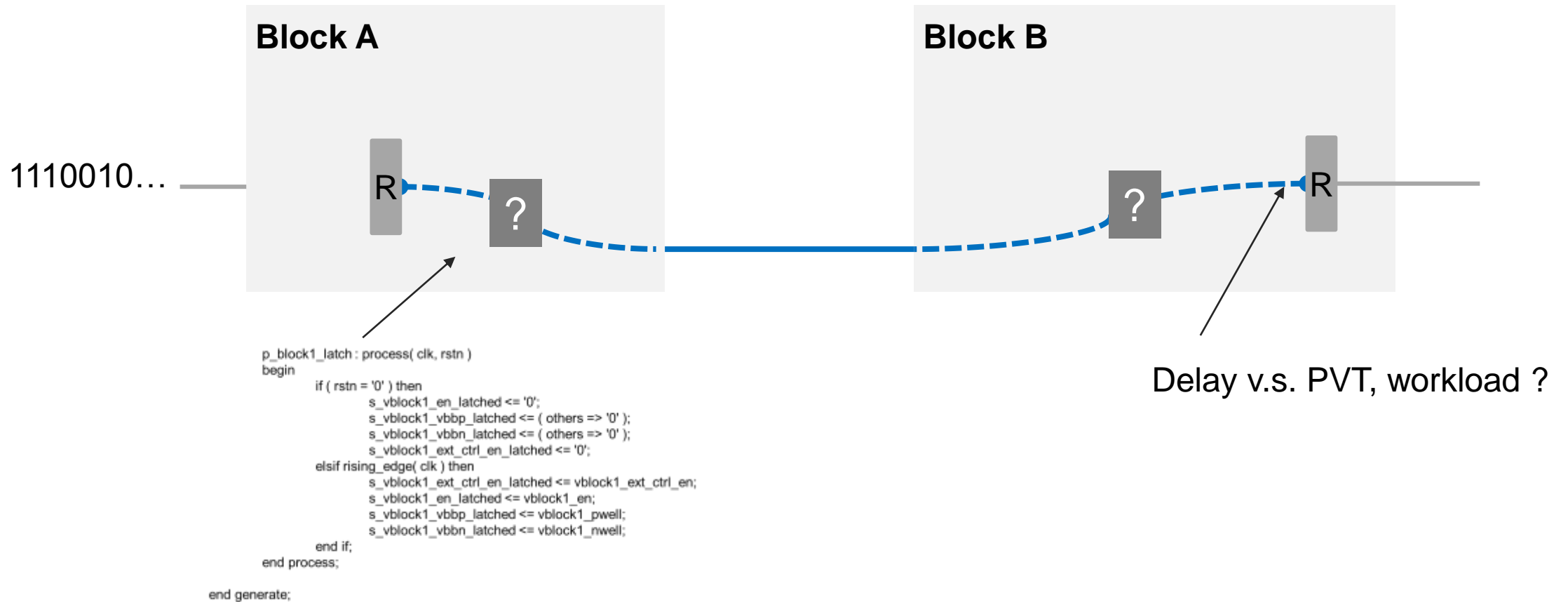
ST: slack time

T_{clk} : clock period

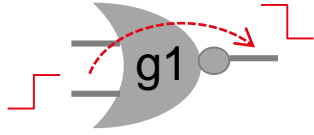
R: register (flip-flop)



RELIABILITY SIMULATION AT REGISTER TRANSFER LEVEL



- Gate degradation model (golden reference)



$$dtg / tg = S * (SP * POT)^n$$

[Huard et al., IRPS]

dtg: delay drift of timing arc

tg: fresh delay

S: sensitivity

n: time exponent

SP: static probability

POT: power-on time

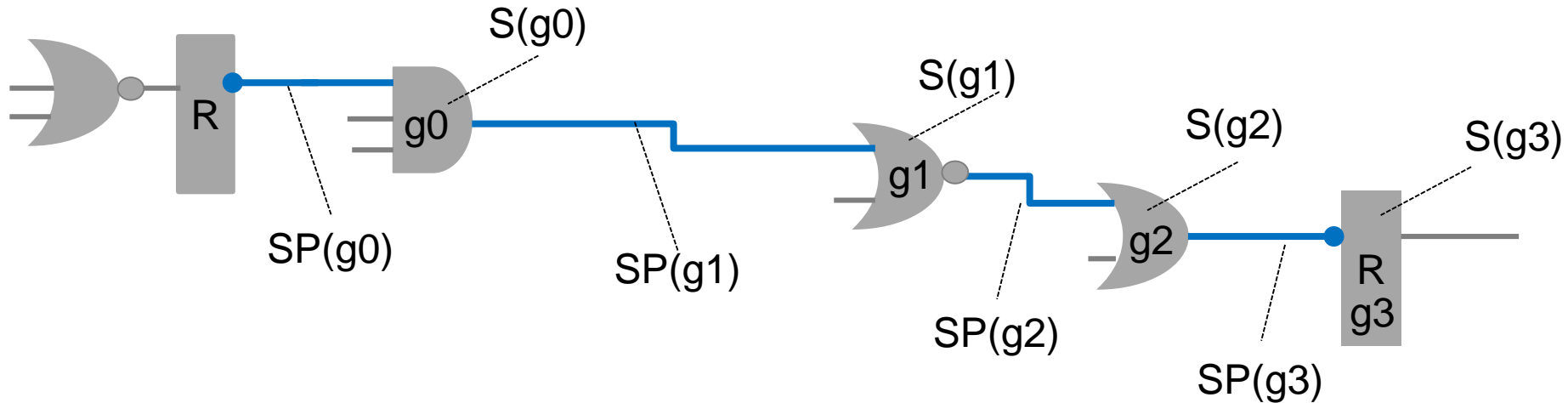
$$SP = \frac{\text{time at 1}}{POT}$$

- Path degradation model at gate level

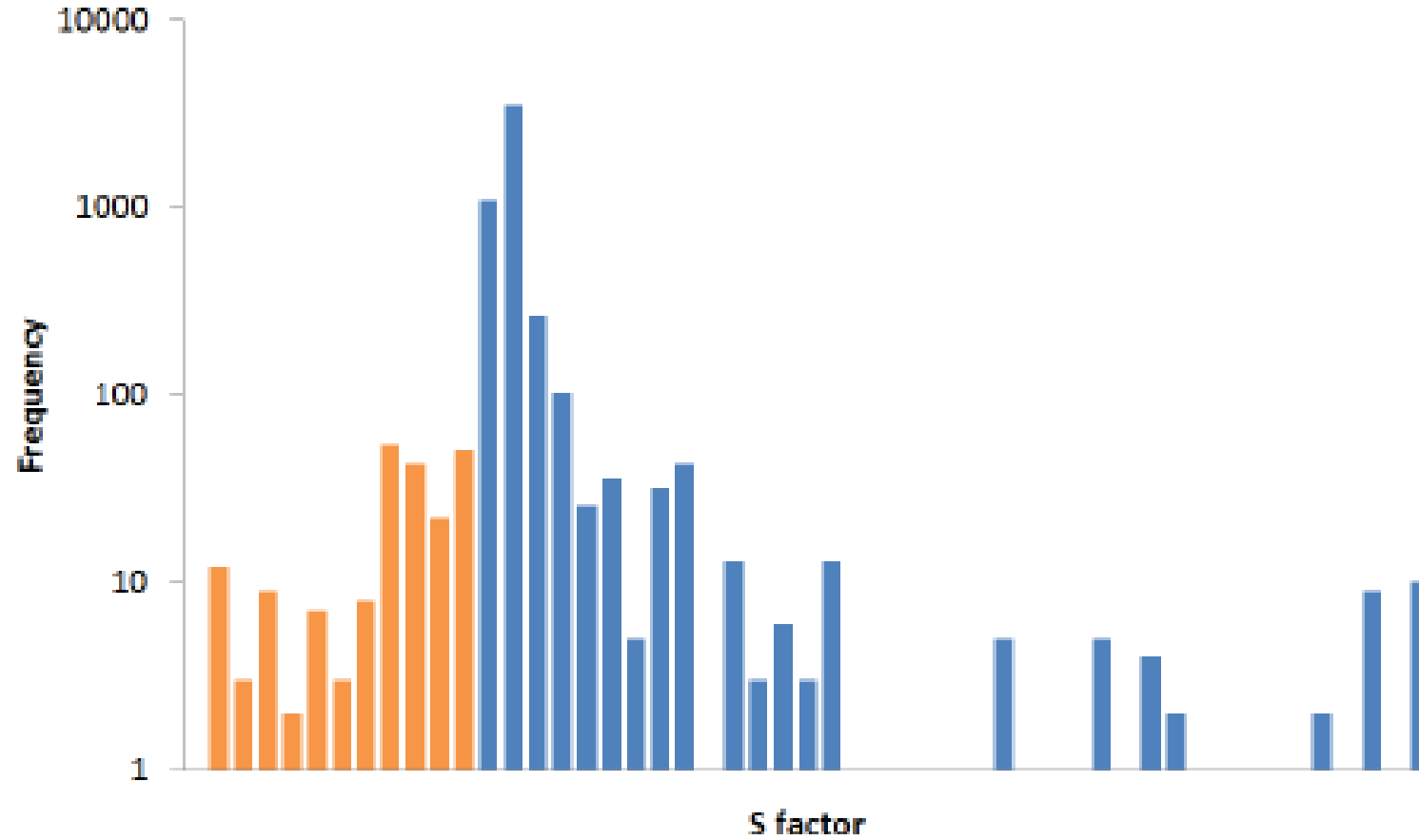
$$dtp / tp = \sum S(g_i) * (SP(g_i) * POT)^n$$

dtp: delay drift of physical path

tp: fresh delay



RELIABILITY SIMULATION AT REGISTER TRANSFER LEVEL

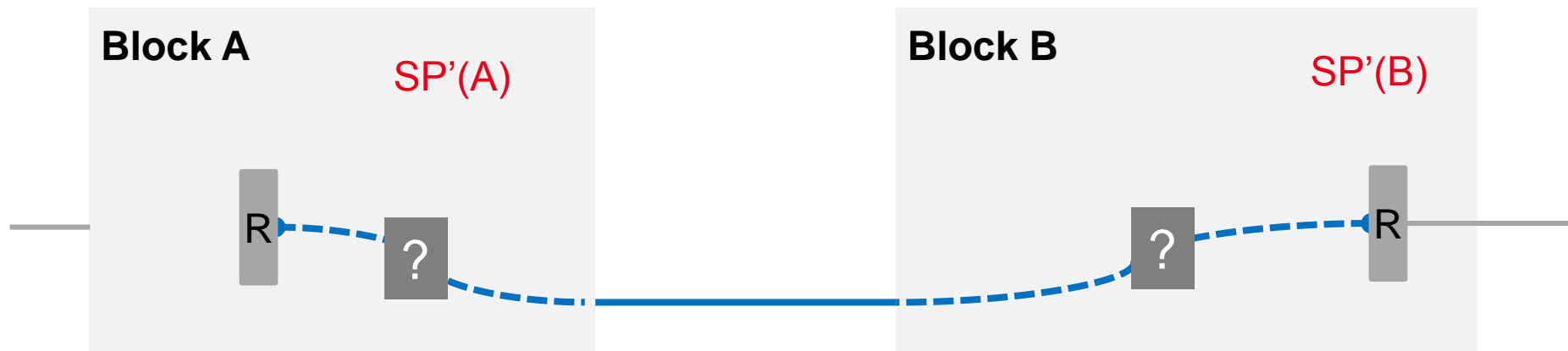


RELIABILITY SIMULATION AT REGISTER TRANSFER LEVEL

- Register-to-Register path model at RTL

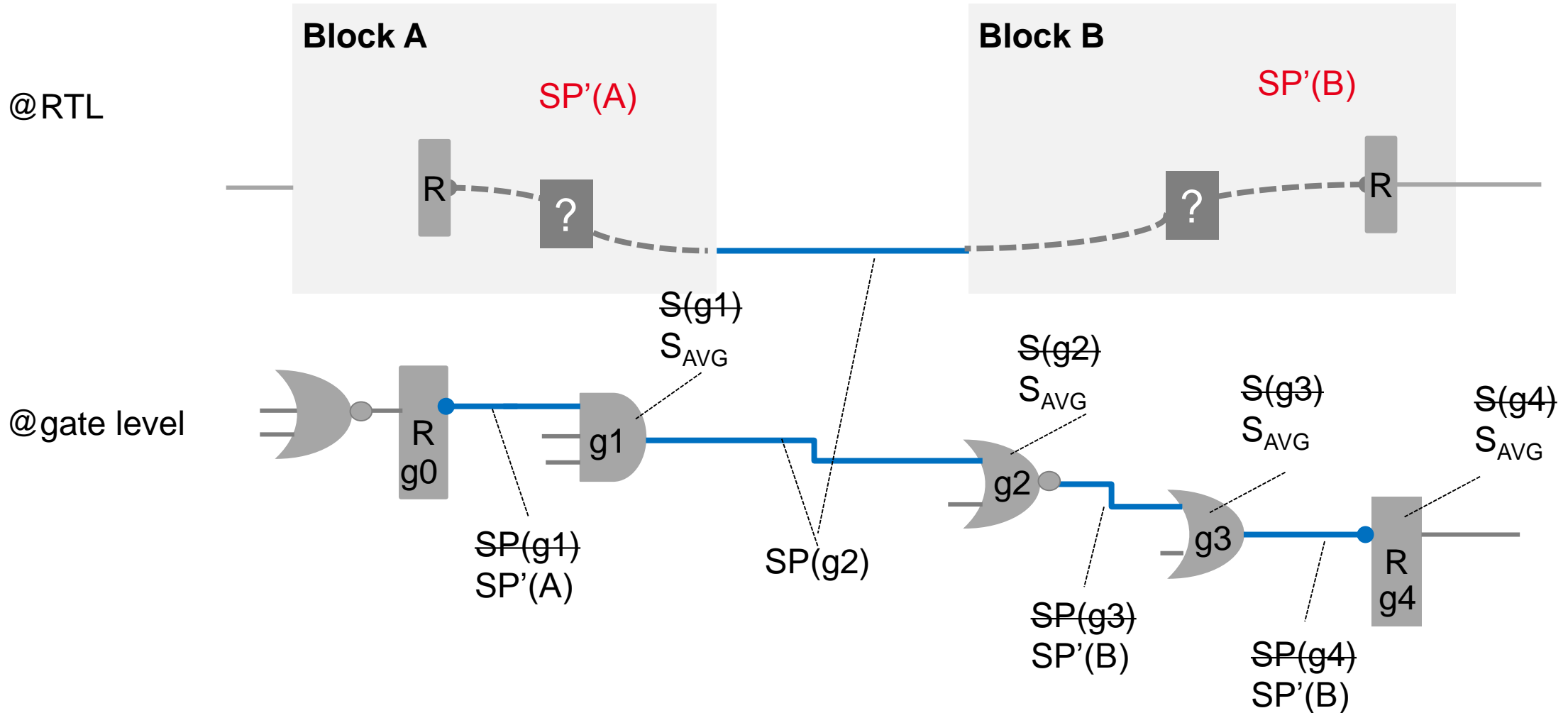
$$dtp / tp \approx S_{AVG} * \sum (SP'(g_i) * POT)^n$$

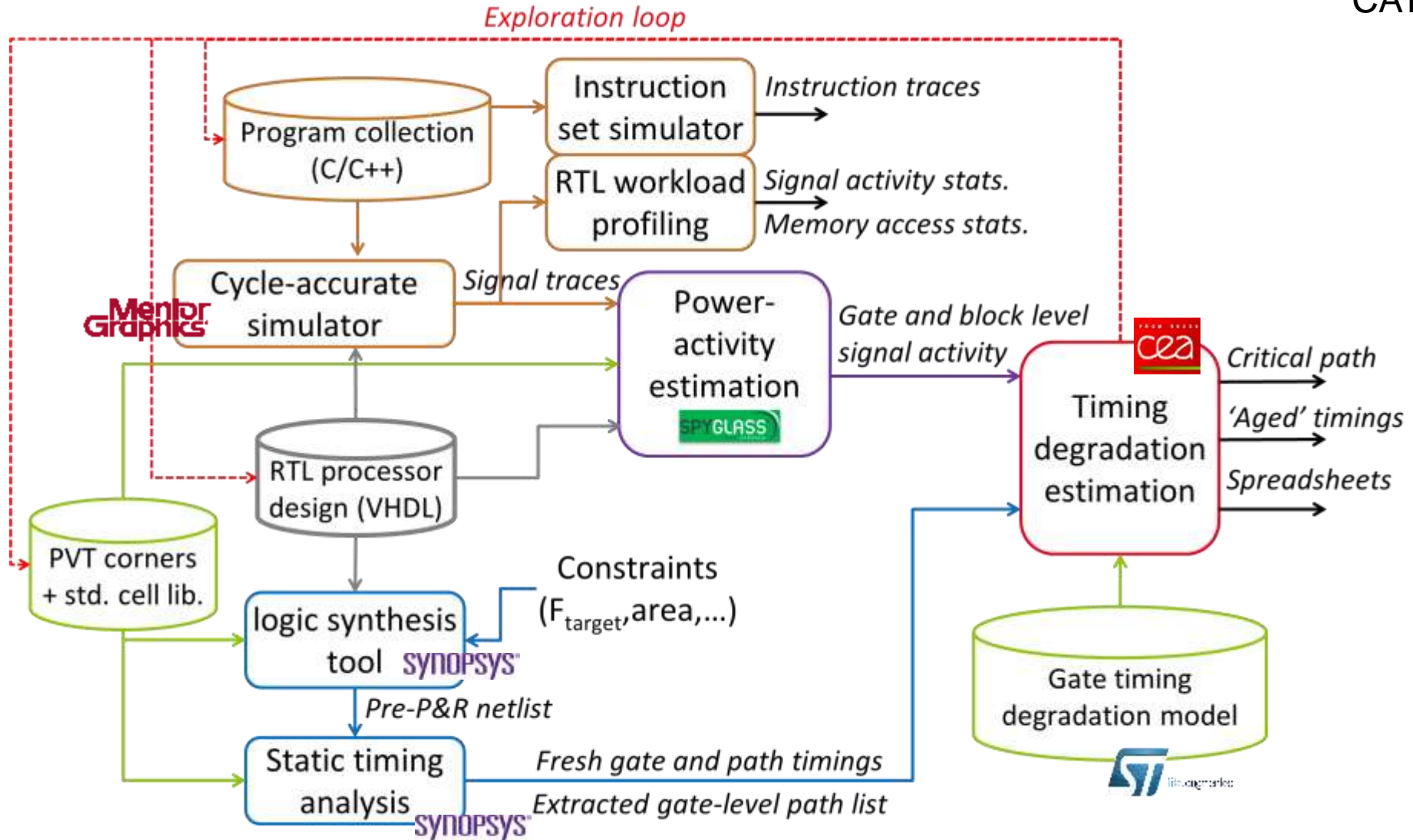
S_{AVG}: Average sensitivity in the whole circuit
SP': approximated SP of "parent" of gate



S_{AVG}

RELIABILITY SIMULATION AT REGISTER TRANSFER LEVEL





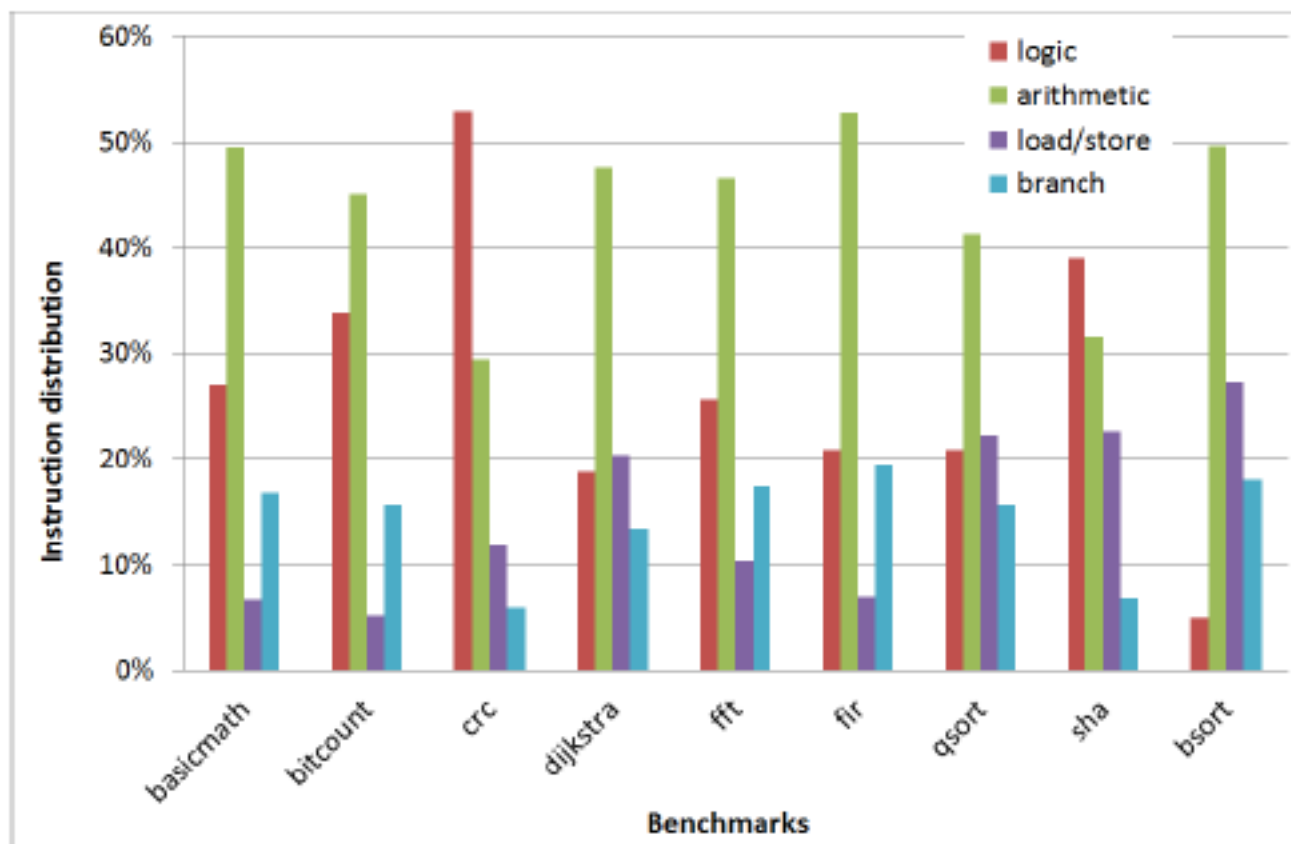
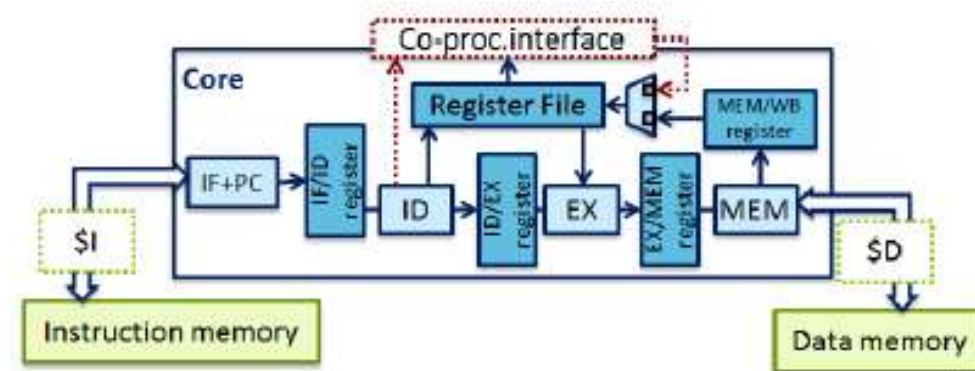
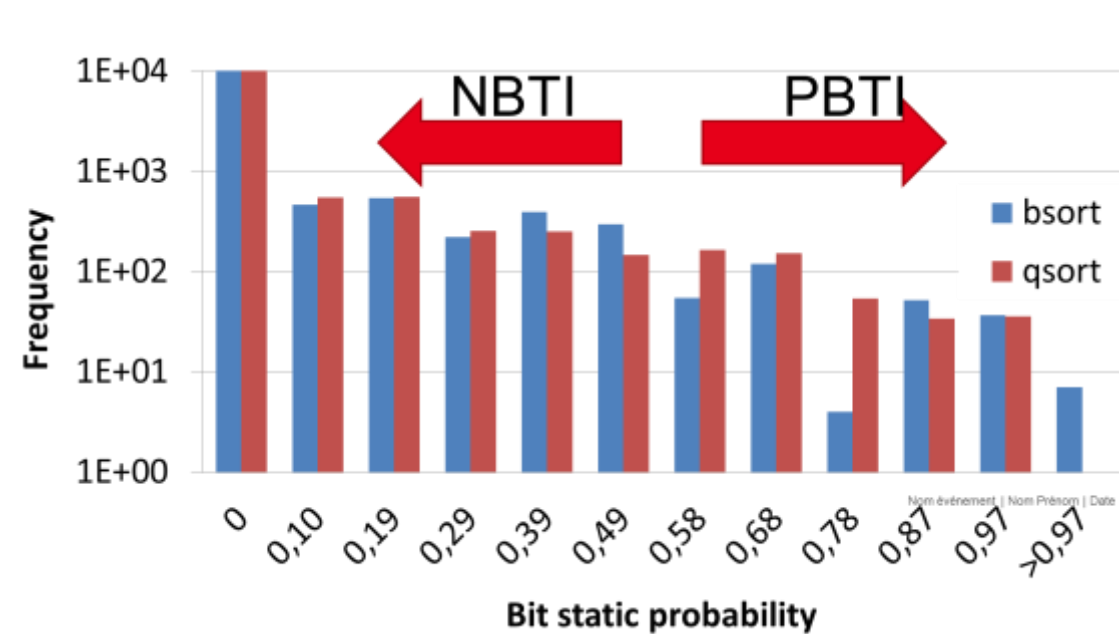


Figure 5. Instruction distribution of benchmarks after simulation with an ISS.

MIPS R4000 like processor (AntX)

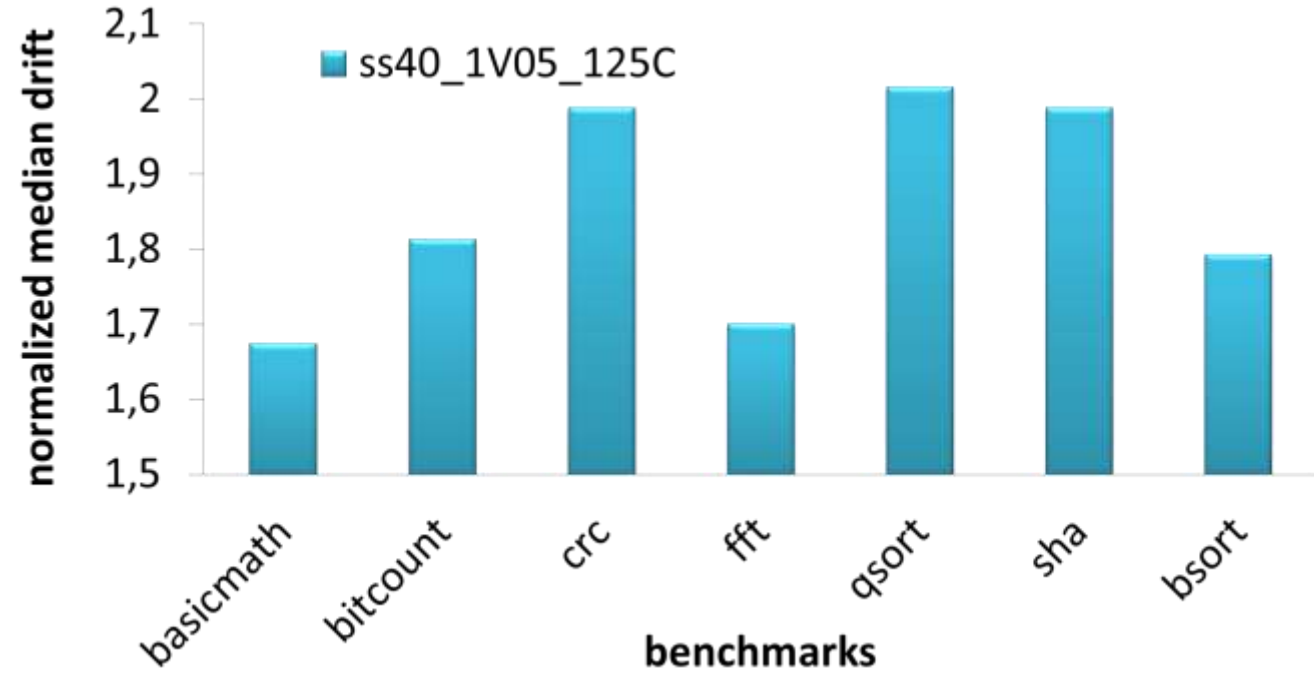


Workload profiling at RTL

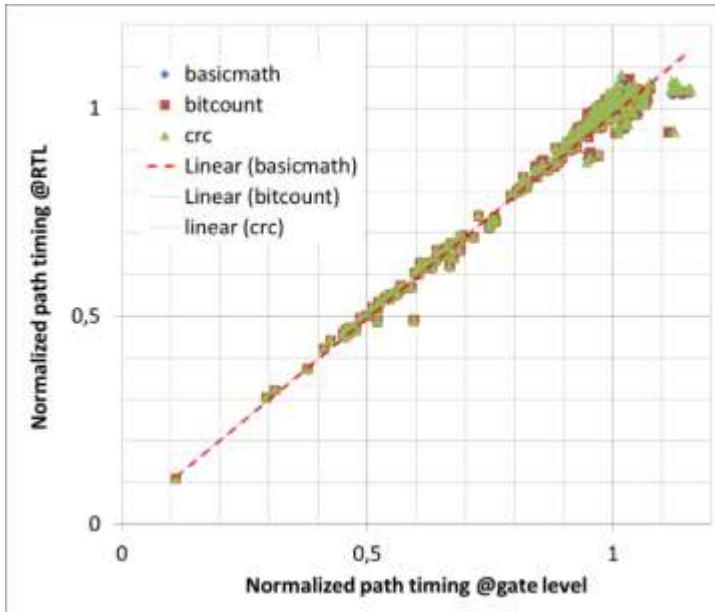


Median drift v.s. benchmarks

ST automotive bulk 40n, 125°C, 1V05 (WC)

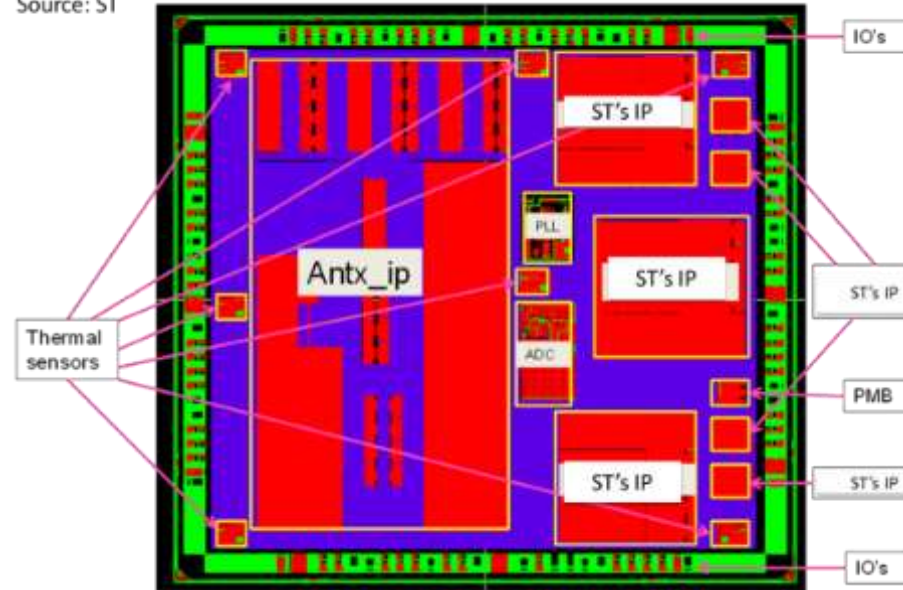


RTL vs. Gate-level (golden)



AntX in SYLVESTERM40 testchip (2015)

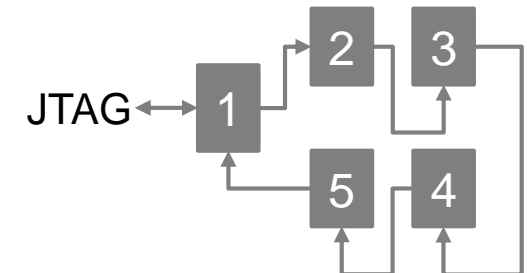
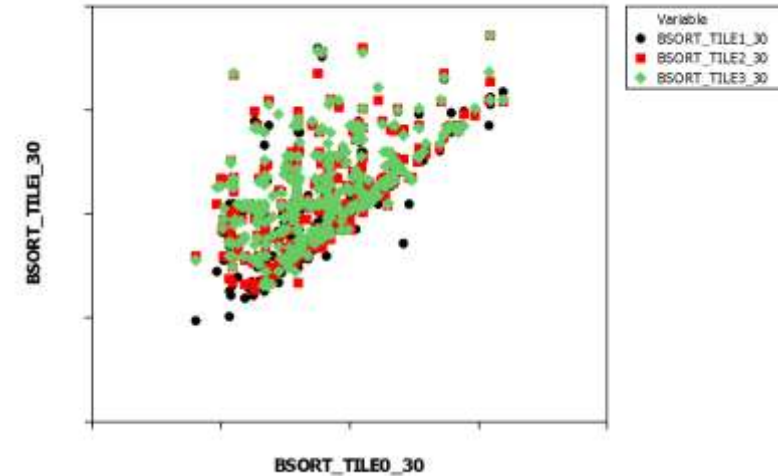
Source: ST



5 AntX processors (CEA),
Power-On SW self-testing,
Self-control program scheduling,
SW checking, ...

COPYRIGHT CEA 2017

VDDmin (Wafer testing)

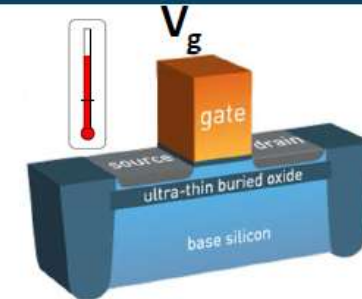


- Voltage
- Temperature
- Workload is not a major driver

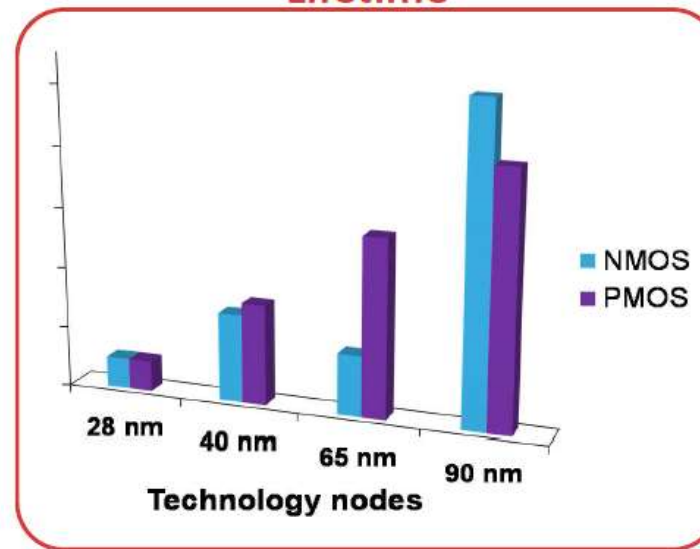
Oxide Breakdown

TDDB degradation

- Dielectric breakdown (TDDB) is enhanced
- EOT reduction:
 - drives the lifetime reduction
 - but makes BD event more progressive

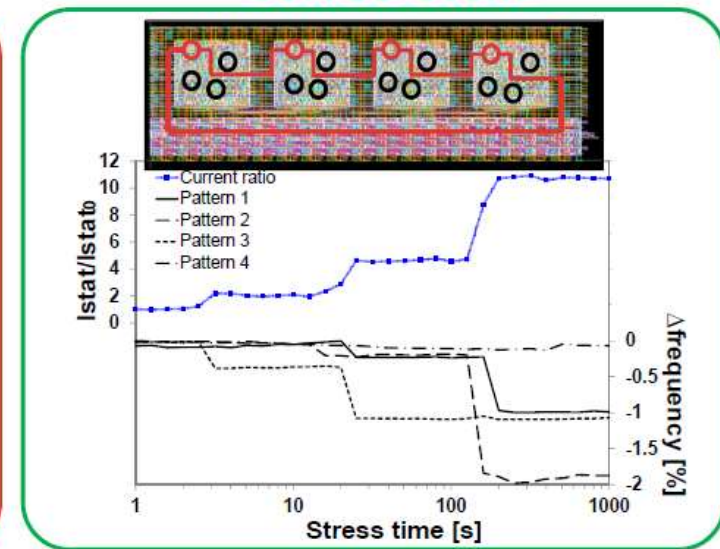


Lifetime



Source : Saliva et al., SETS, 2015
Source : Saliva et al., IRPS, 2014

Soft BD event

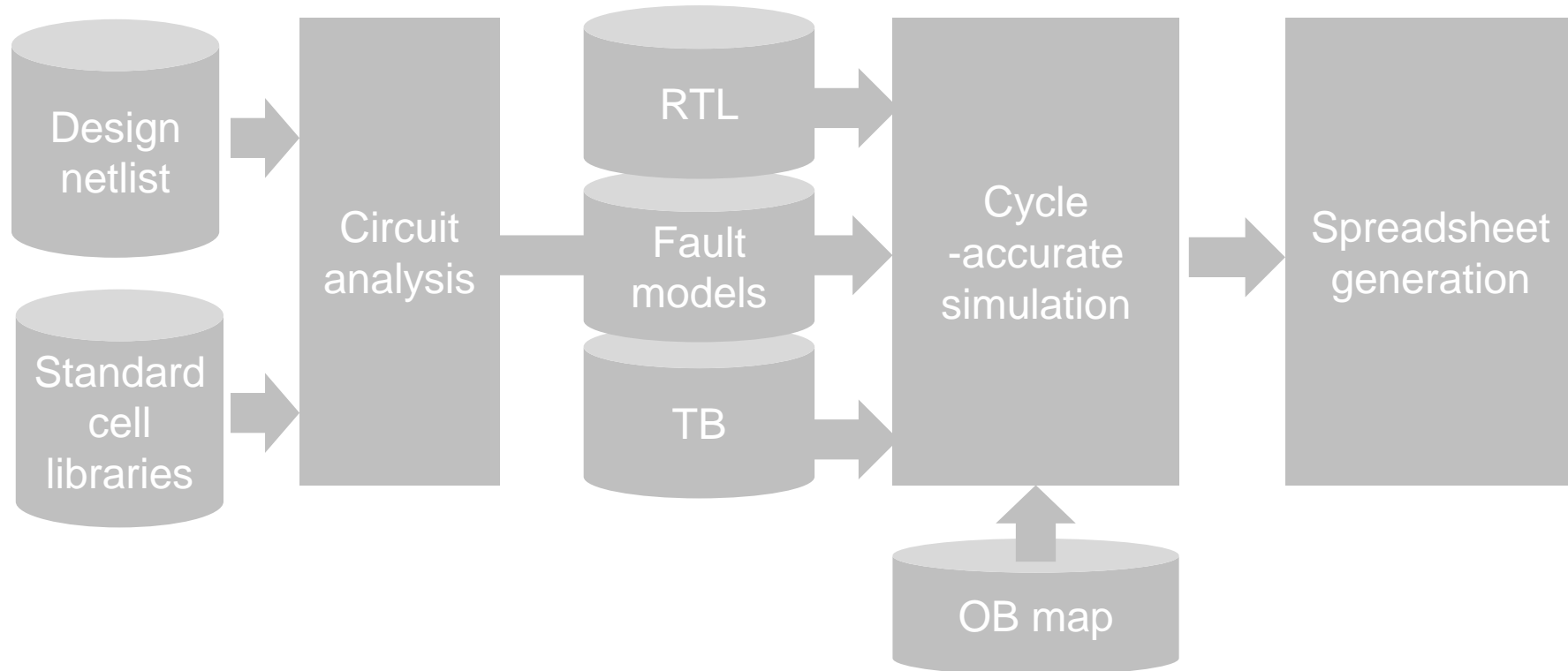
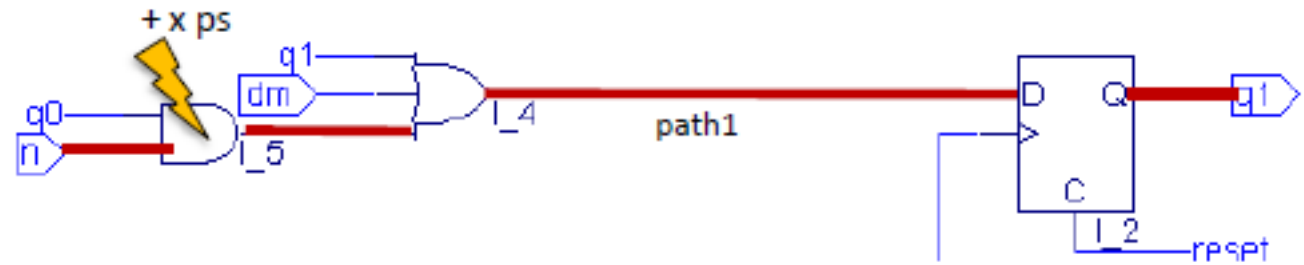


Vincent Huard / STMicroelectronics

13

OXIDE BREAKDOWN ASSESSMENT

- Assessment at RTL



OXIDE BREAKDOWN ASSESSMENT

- AntX processor
- Fault injection on 180 flip-flops => single-fault propagation probability
 - Flip-flops belonging to the 100000 paths (over 2000000) with shortest slack time
- 361 simulations

memtest86

	pmos	nmos	tot
No errors	71	71	142
Silent errors	36	25	61
Single error	0	1	1
Multiple errors	73	83	156

bitcount

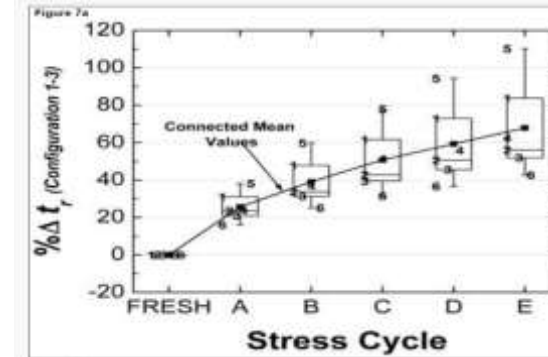
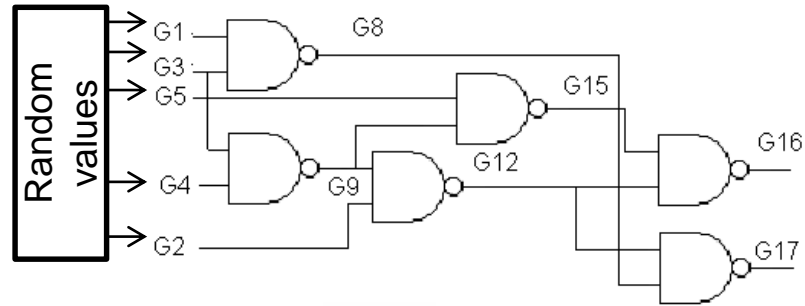
	pmos	nmos	tot
No errors	63	63	126
Silent errors	56	73	129
Single error	1	0	1
Multiple errors	60	44	104

MODELLING BY STOCHASTIC PROCESS: OXIDE BREAKDOWN

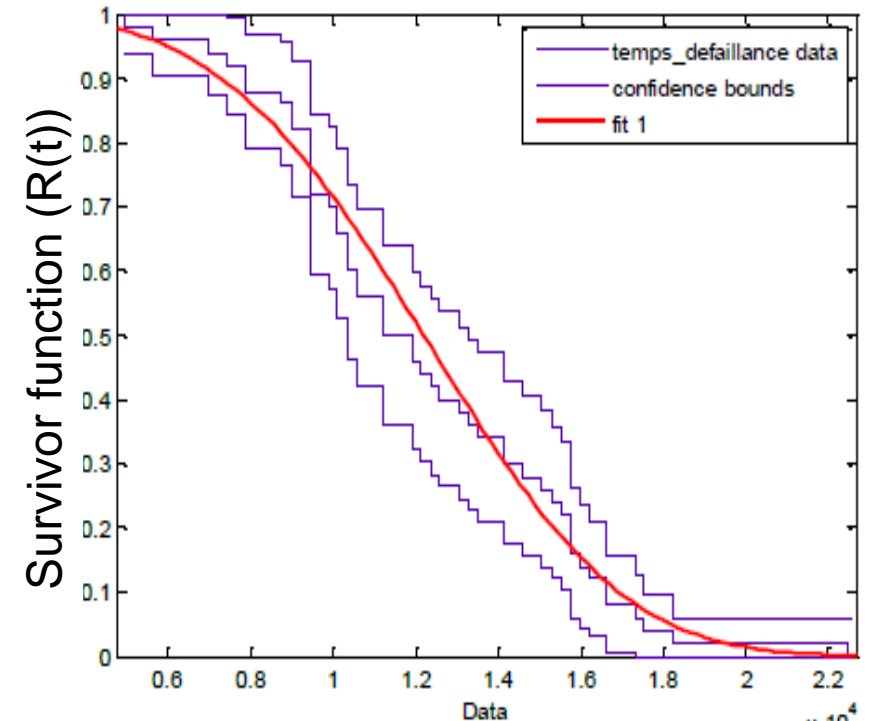
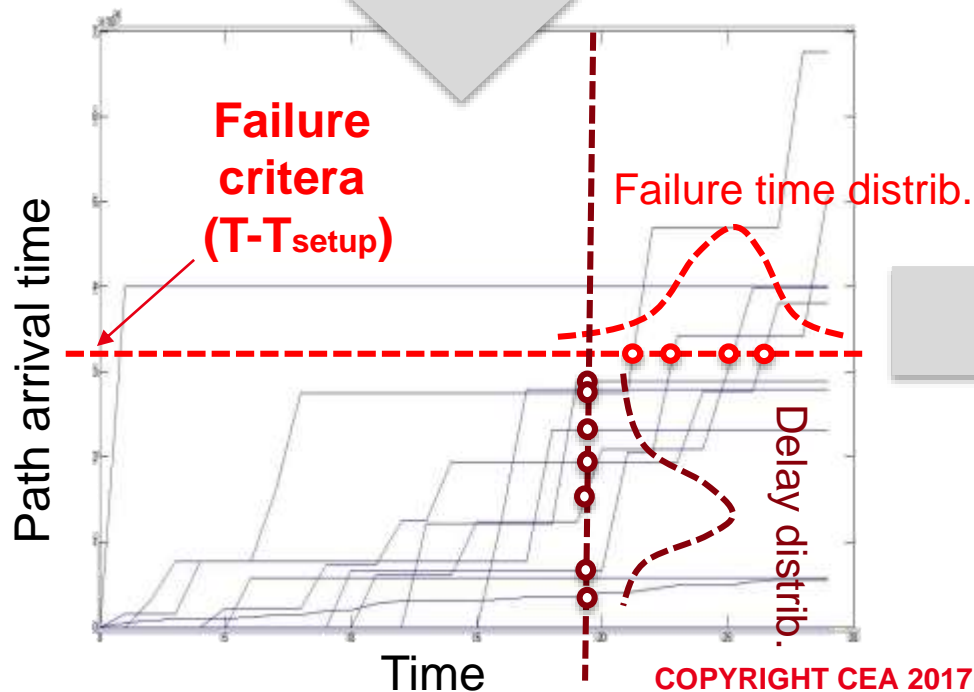
[Heron, Guérin et al., ALT'12]

Gamma's stochastic process: $\% \Delta d = \gamma(y(d_i) - y(d_{i-1}), \beta)$

(C17 - ISCAS85)

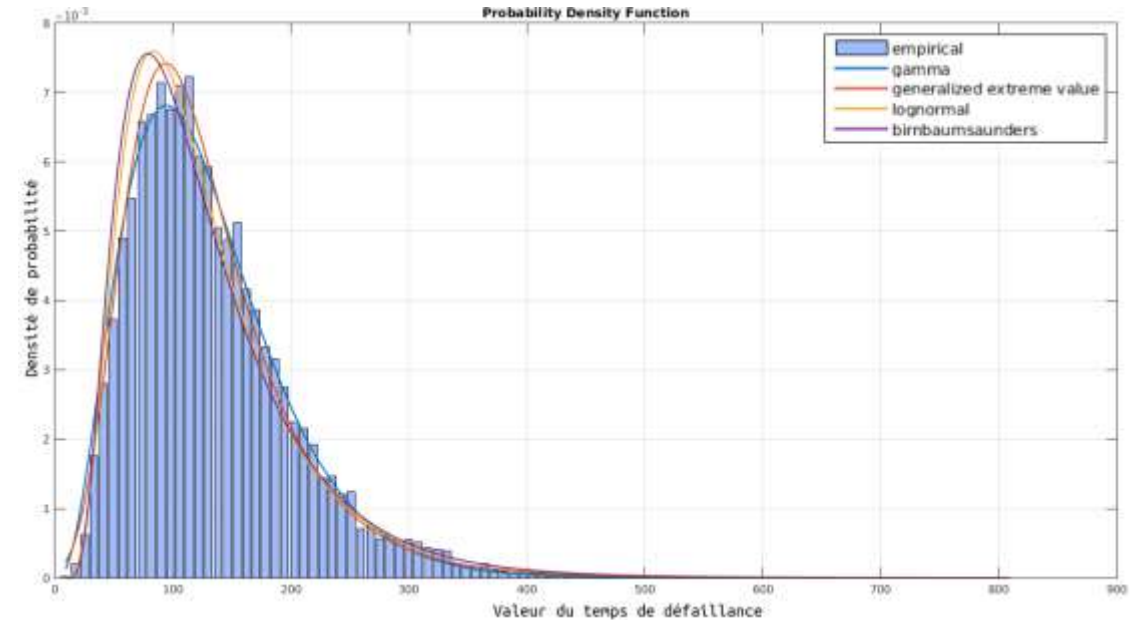
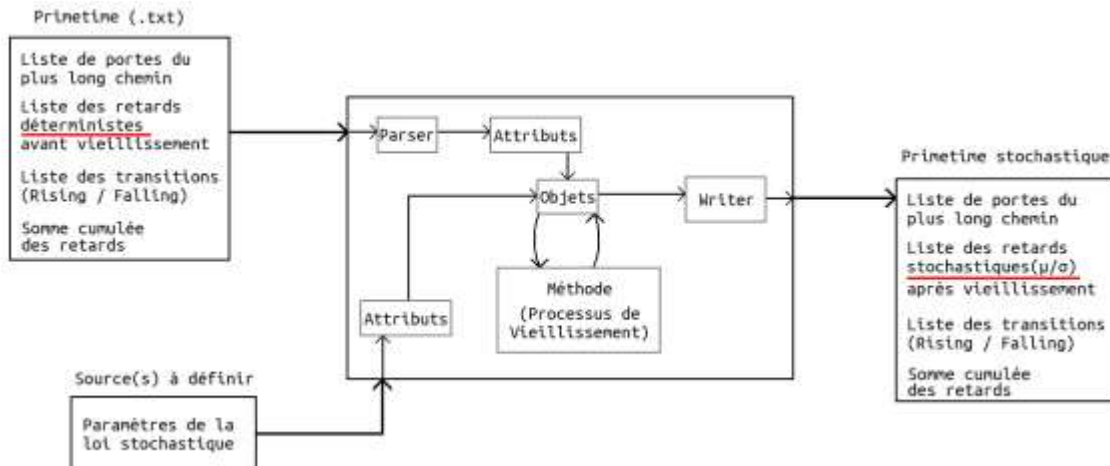
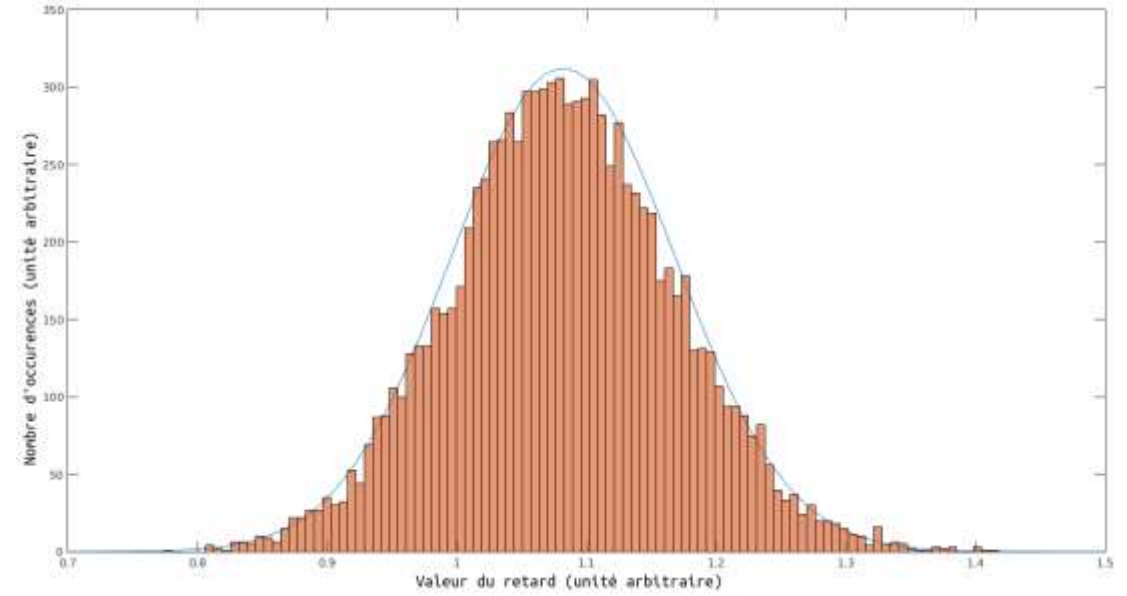
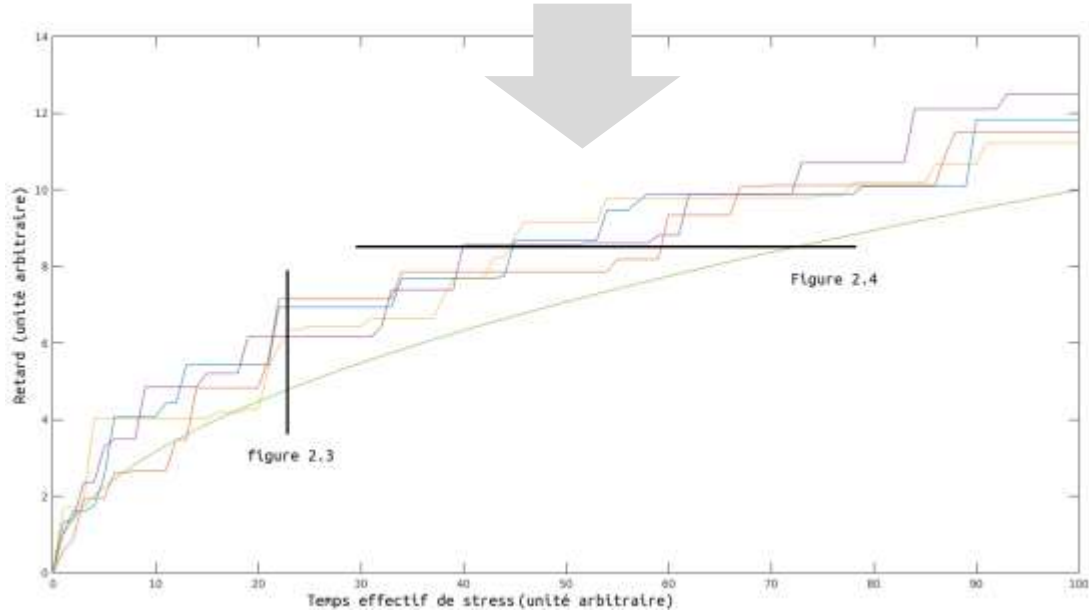


Monte Carlo based simulation (Matlab)



MODELLING BY STOCHASTIC PROCESS: BTI

J.Fang et al Understanding the impact of transistor-level bti variability. IRPS 2012





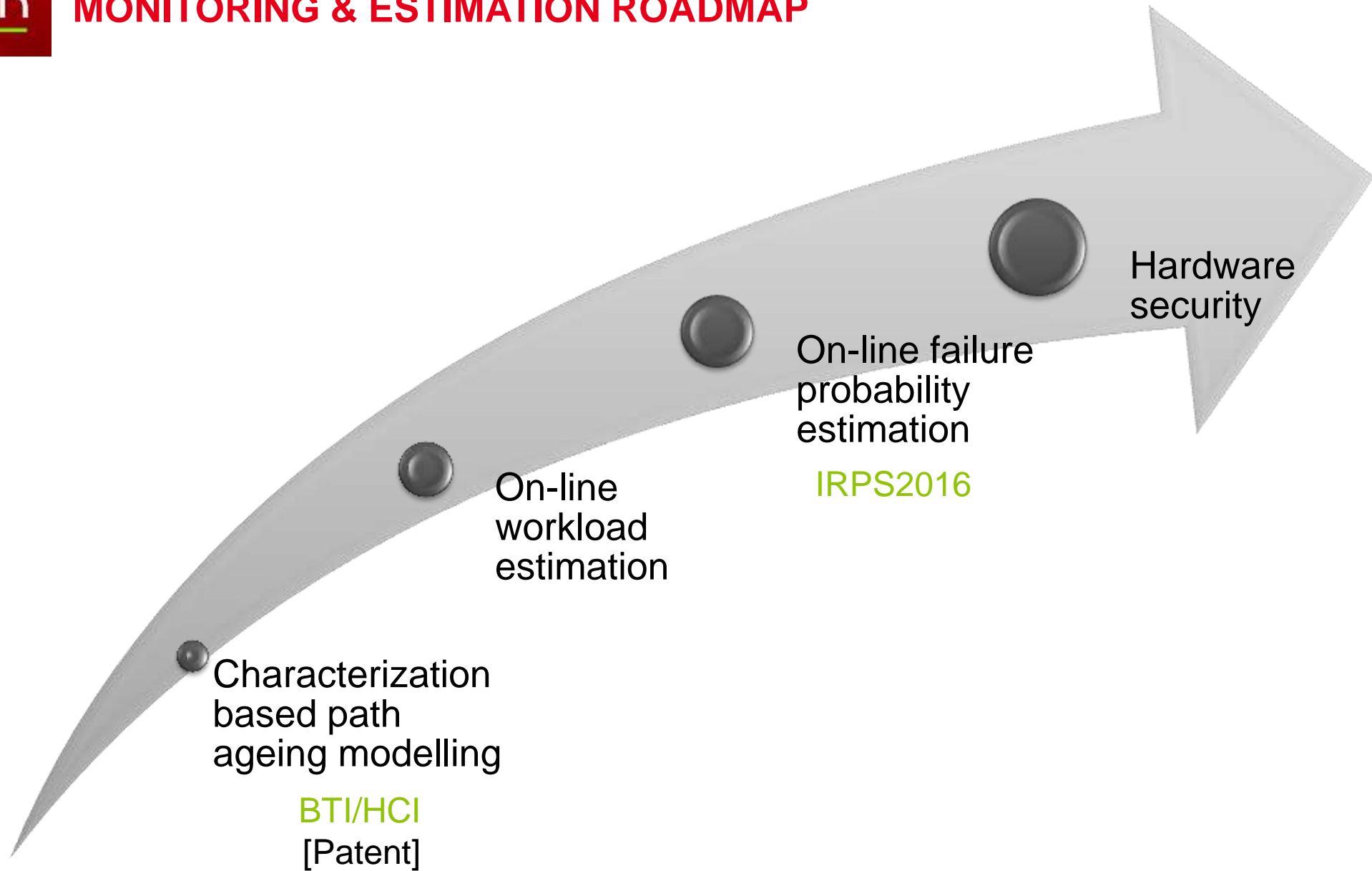
PART IV: AGEING MONITORING & ESTIMATION

How to build accurate circuit ageing models independently of specific device models?

How to derive accurate ageing information from PVT sensors?

How to estimate actual workload for ageing estimation?

MONITORING & ESTIMATION ROADMAP



ON-CHIP MONITORING FOR AGEING CHARACTERIZATION OF DEVICES

(19) **United States**

(12) **Patent Application Publication**
Mikkola

(10) **Pub. No.:** US 2012/0245879 A1

(43) **Pub. Date:** Sep. 27, 2012

(54) **PROGRAMMABLE TEST CHIP, SYSTEM AND METHOD FOR CHARACTERIZATION OF INTEGRATED CIRCUIT FABRICATION PROCESSES**

(75) **Inventor:** Esko O. Mikkola, Tucson, AZ (US)

(73) **Assignee:** Ridgetop Group, Inc.

(21) **Appl. No.:** 13/424,025

(22) **Filed:** Mar. 19, 2012

Related U.S. Application Data

(60) Provisional application No. 61/465,463, filed on Mar. 21, 2011.

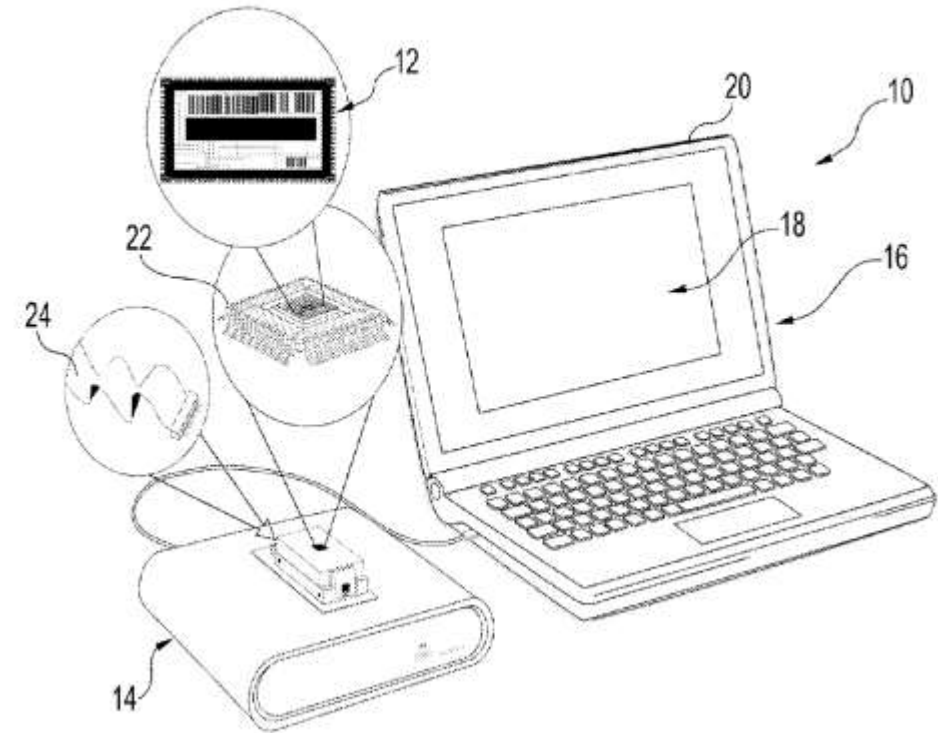
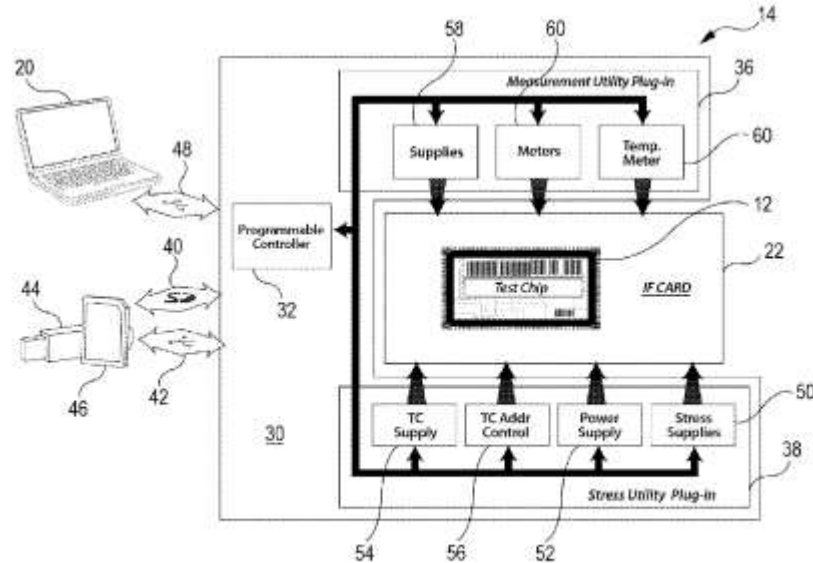
Publication Classification

(51) **Int. Cl.**
G06F 19/00 (2011.01)

(52) **U.S. Cl.** 702/117

(57) **ABSTRACT**

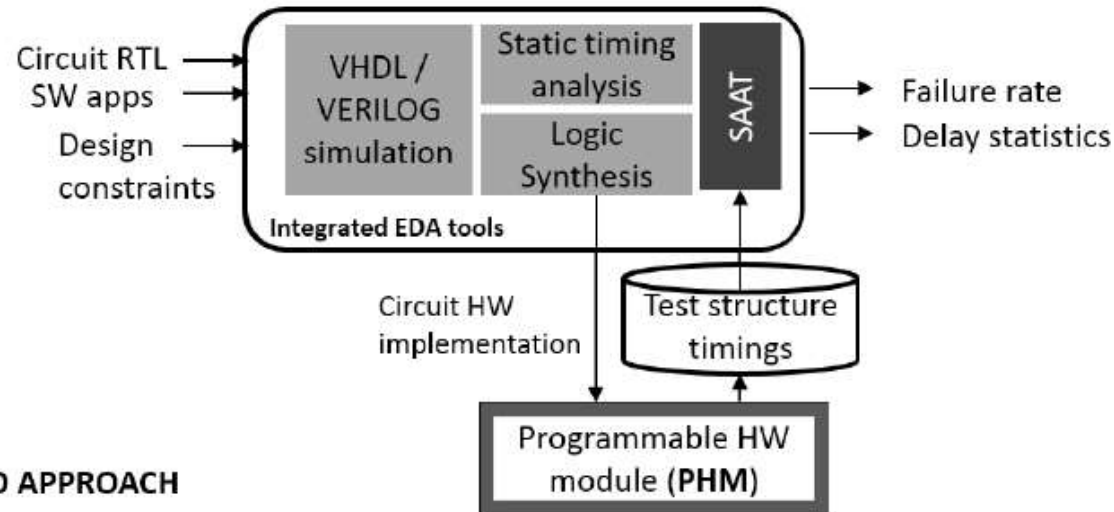
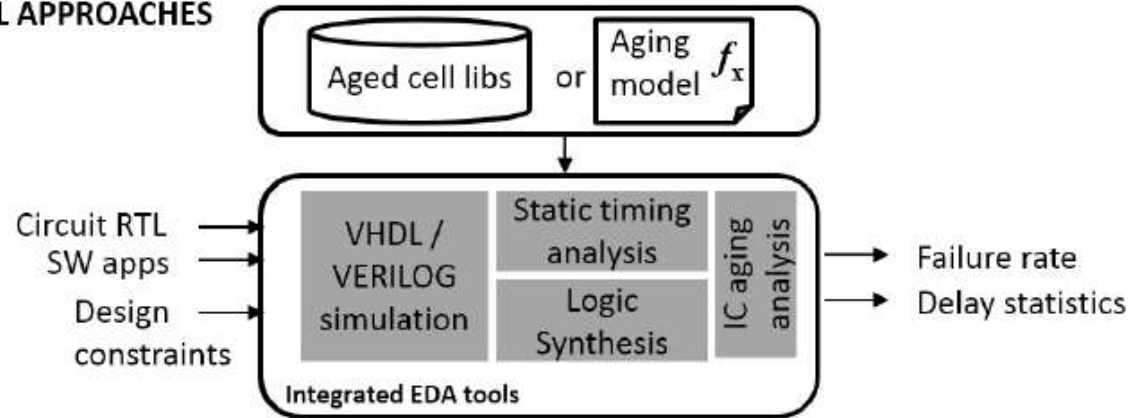
A test chip, system and method for testing large numbers of test devices on a single test chip decreases the time and complexity required to characterize the variation and reliability of the IC fabrication process. A remotely configurable test chip can be programmed with varying bias conditions for testing of process variation or numerous failure modes on large sample sizes. An on-chip addressing technique allows large numbers of test devices to be tested simultaneously and the measurement signals read out serially for different test devices. The test chip may be configured for wafer, die or package-level testing.



ON-CHIP MONITORING FOR AGEING CHARACTERIZATION OF CIRCUIT

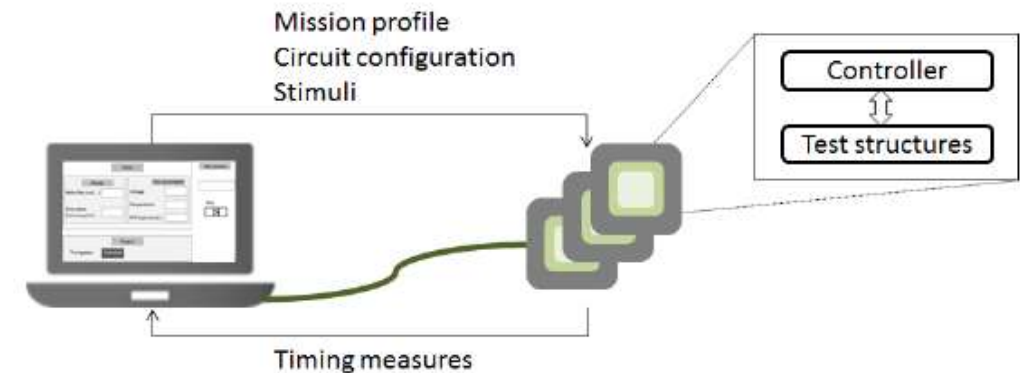
- Ready to be inserted in a CAD flow (RTL)
- Not for sign-off but for early design decisions
- No need for modelling of device physics of failure

CLASSICAL APPROACHES



PROPOSED APPROACH

SAAT: Software Aging Analysis Tool



- Preliminary results on AntX processor
- FDSOI28n ($V_b=0$)
- Critical path analysis (Fresh frequency 500MHz@0V6-125C)

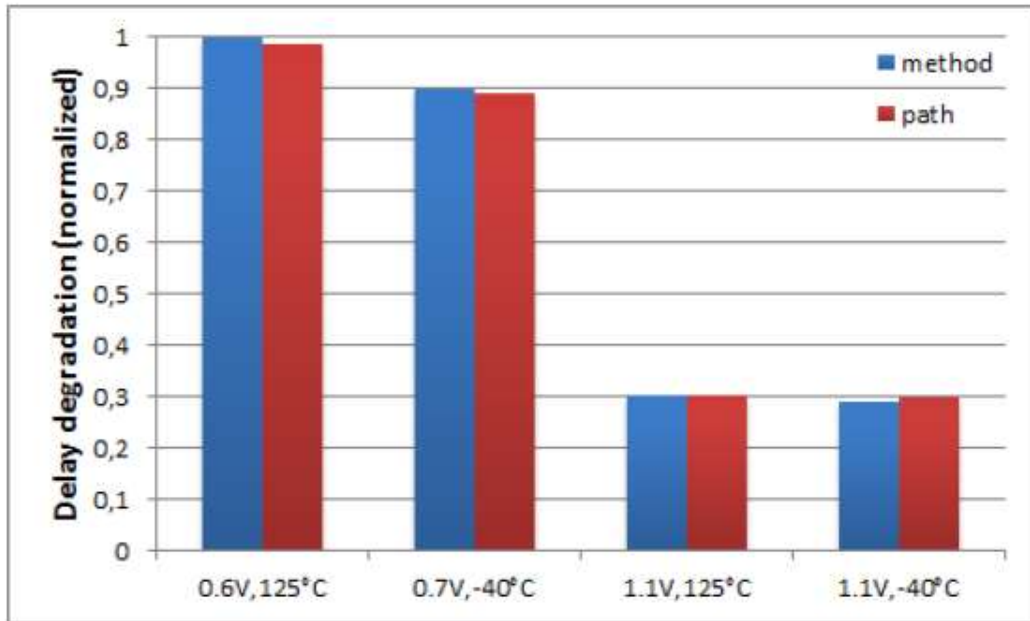


Fig. 6. Accuracy between the proposed method and the observation of the delay degradation on the path with ΔV_{th} constant

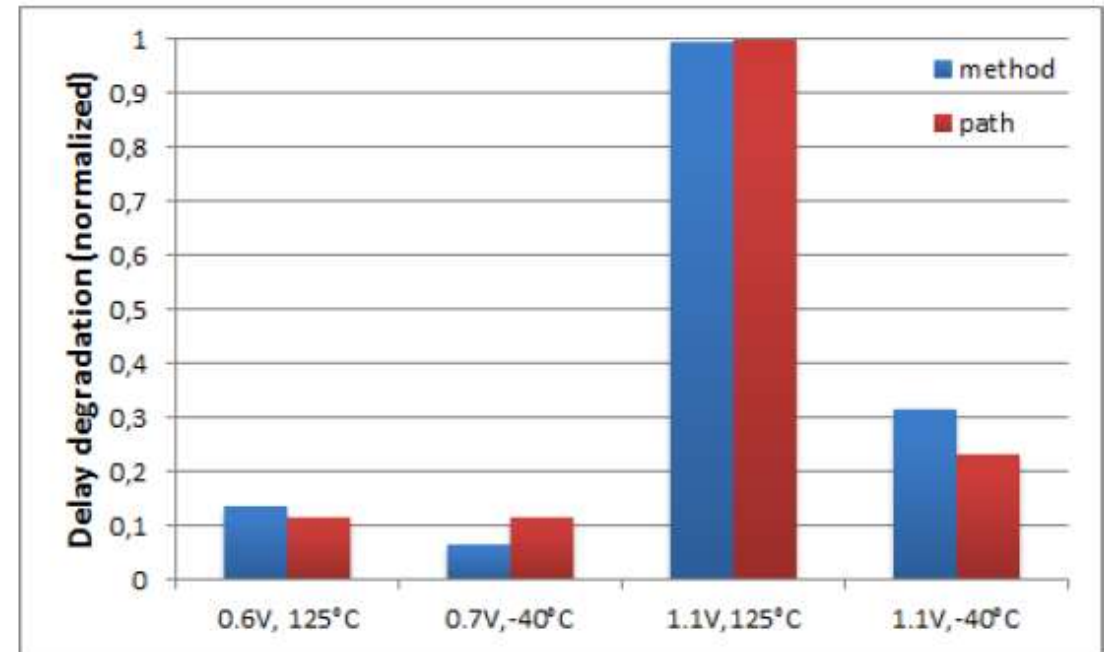
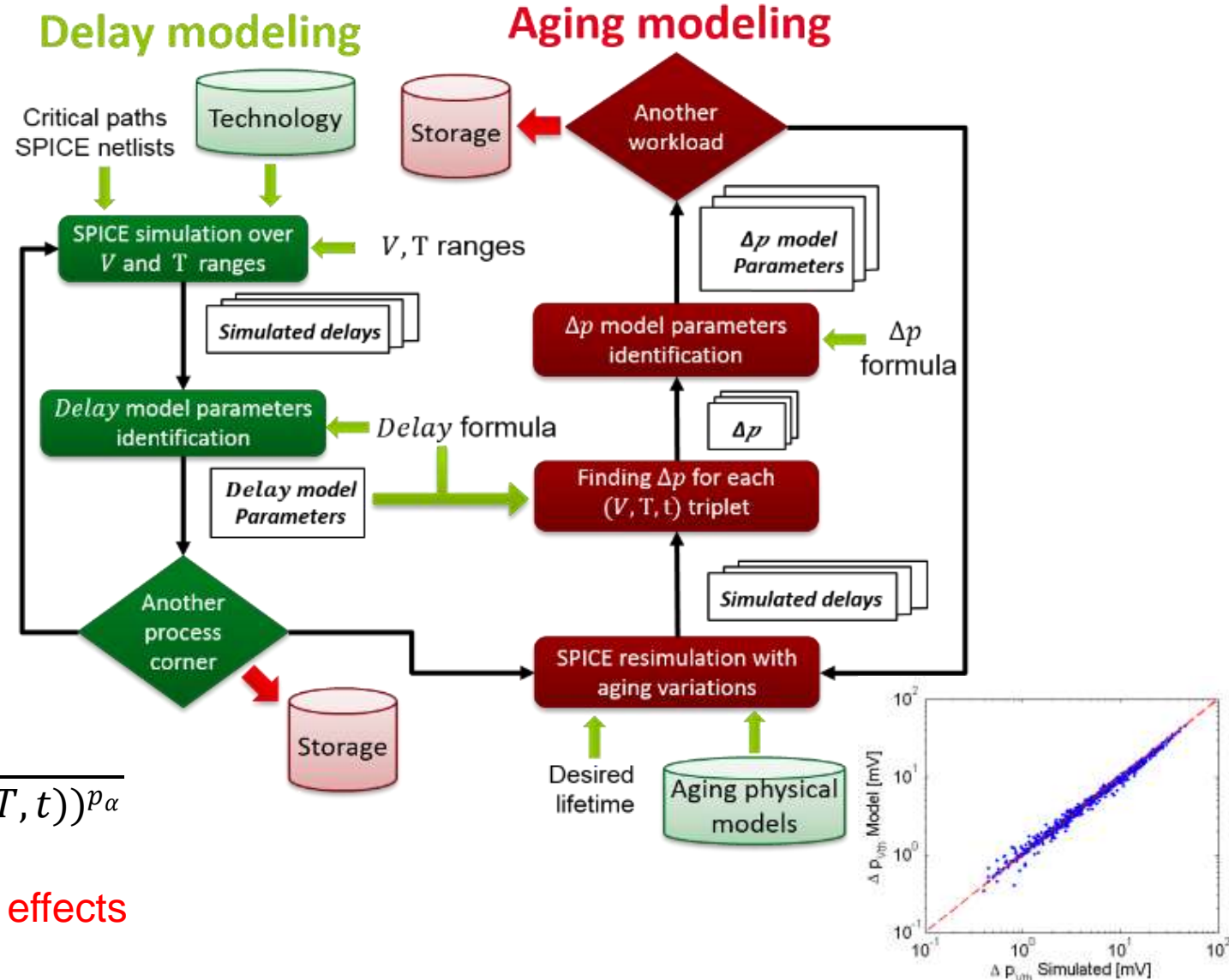


Fig. 7. Accuracy between the proposed method and the observation of the delay degradation on the path with ΔV_{th} variable

- Mechanism to on-line assess the circuit reliability by estimating the degradation of its critical paths under the actual stress conditions.
- Creating accurate but nonetheless simplified circuit-level aging models from existing device-level models and using in-situ monitors to follow dynamic (slow) variations



$$Delay(V, T, t) = p_{\beta} + \frac{p_{\mu}^{-1}(T) * V}{(V - (p_{V_{th}}(T) + \Delta p_{V_{th}}(V, T, t)))^{p_{\alpha}}}$$

Critical path delay

Ageing effects

- **Development time and cost**
 - Early analysis → re-design effort is reduced
- **Engineer productivity**
 - Models aid to predict the parameter shift per device
- **Risk management**
 - Design risk issues can be evaluated earlier in the design flow
- **Repeatability**
 - A change in the design specification can be taken into account immediately
- **Technology versatility**
 - TSMC 45n, ST bulk40n, ST FDSOI28n and others
 - HW-assisted ageing modelling
- **Market segments**
 - Safety critical applications: small/medium size ICs (e.g. microcontroller)
 - Telecom, wireless, consumer: High-end ICs (e.g. multicore)

AGEING INDUCED SECURITY THREATS

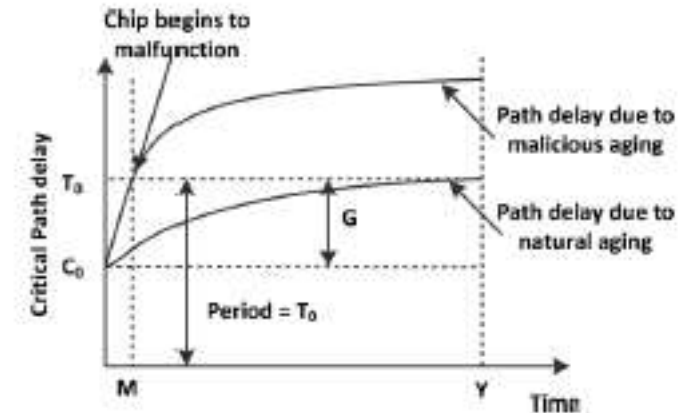
MAGIC: Malicious Aging in Circuits/Cores [Karimi et al, TACO 2015]




- **Attaque qui accélère les effets du NBTI**
 - Erreurs observées après 1 mois
- **Processeur OpenSPARC T1, technologie 45 nm**
- **Techniques de mitigation et correction existantes pas suffisantes**

Temps	Dégradation avec attaque	Dégradation nominal
1 mois	10.92%	0.28%
2 mois	13.25%	Pas communiqué
6 mois	16.8%	0.41%

Guardband 10%



AGEING INDUCED SECURITY THREATS

- **Scenario 1 : Attaque pour garantie**
 - Echange du dispositif avant la fin de la garantie
 - **Scenario 2 : Obsolescence programmée**
 - Programme de vieillissement envoyé par le fabricant comme mise à jour
 - **Scenario 3 : Hardware backdoor**
 - Dispositif mis hors service à distance
 - **Scenario 4 : Attaque contre la sécurité**
 - Révélation de clés
- 
- **Comment distinguer le vrai vieillissement du vieillissement par attaque ?**



SOME PERSPECTIVES...

■ Error & performance degradation flags

