Lifetime Reliability Aware Resource Management and Computing: Challenges & Opportunities

Sorin Cotofana

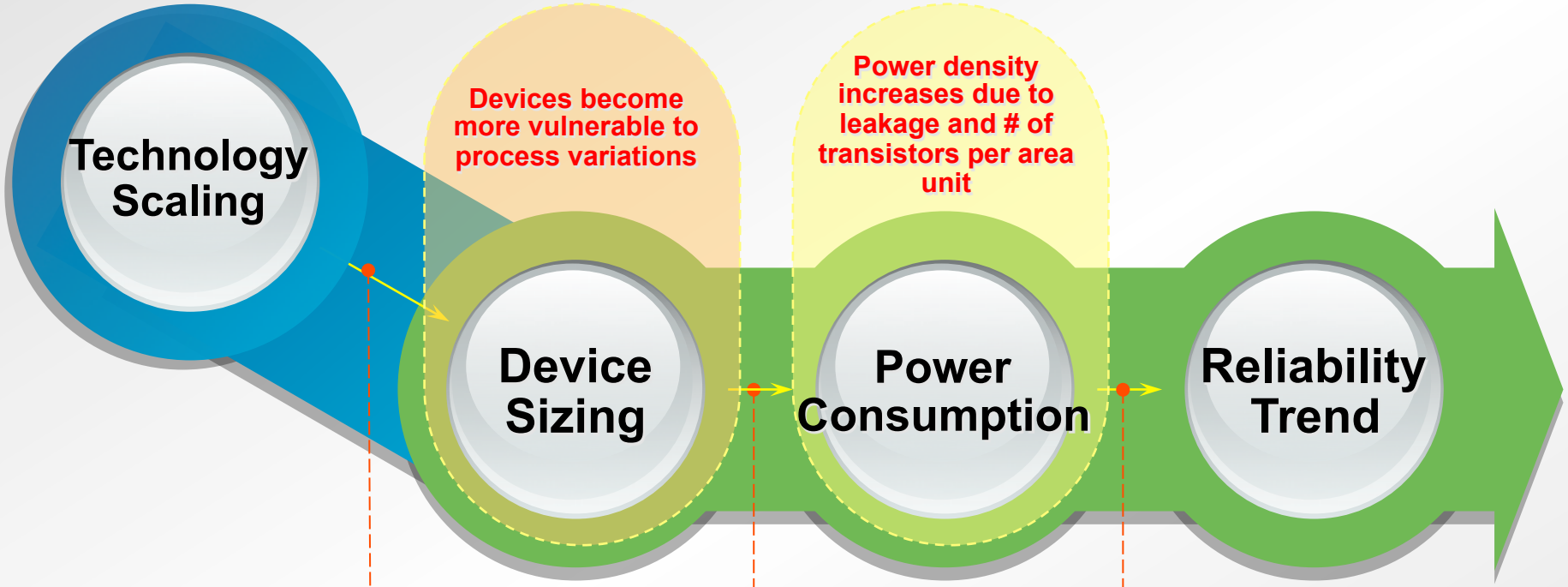
Computer Engineering Laboratory
Delft University of Technology

TU Delft

Contents

TU Delft

Why Reliability Becomes an Issue?



Non-Ideal Supply Voltage Scaling.

- E_{ox} ↑
- Leakage ↑
- # of Transistors ↑

- Power Density ↑
- On Chip Temperature ↑
- Aging ↑

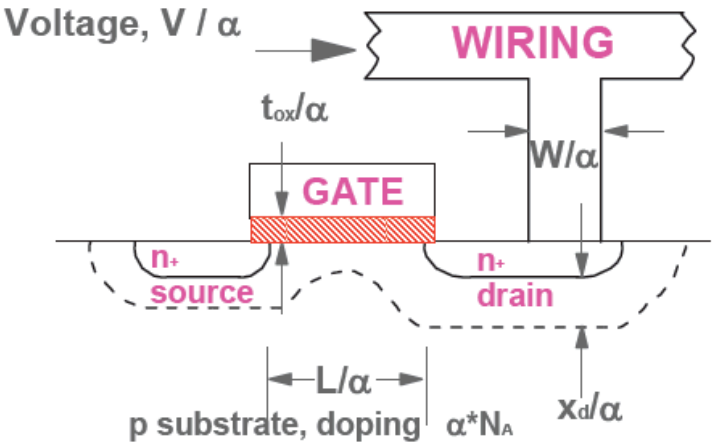
- Reliability per Device ↓
- Reliability per Chip ↓

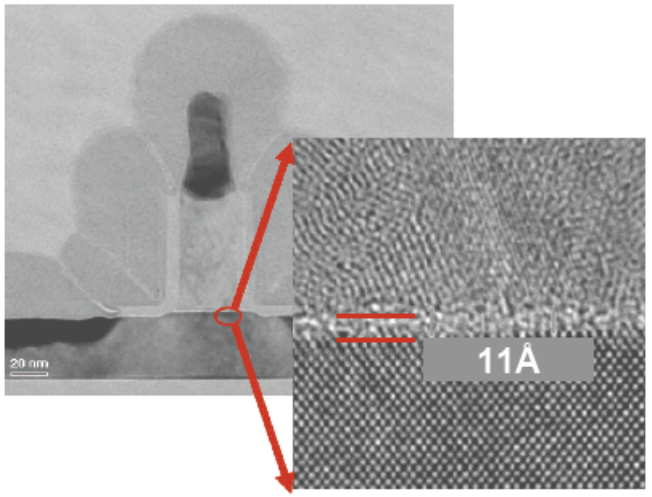
TU Delft

Device Scaling

Non-Ideal Scaling

CMOS Scaling Rules





SCALING:

Voltage:	V/α
Oxide:	t_{ox}/α
Wire width:	W/α
Gate width:	L/α
Diffusion:	x_d/α
Substrate:	$\alpha * N_A$

R. H. Dennard *et al.*,
IEEE J. Solid State Circuits, (1974).

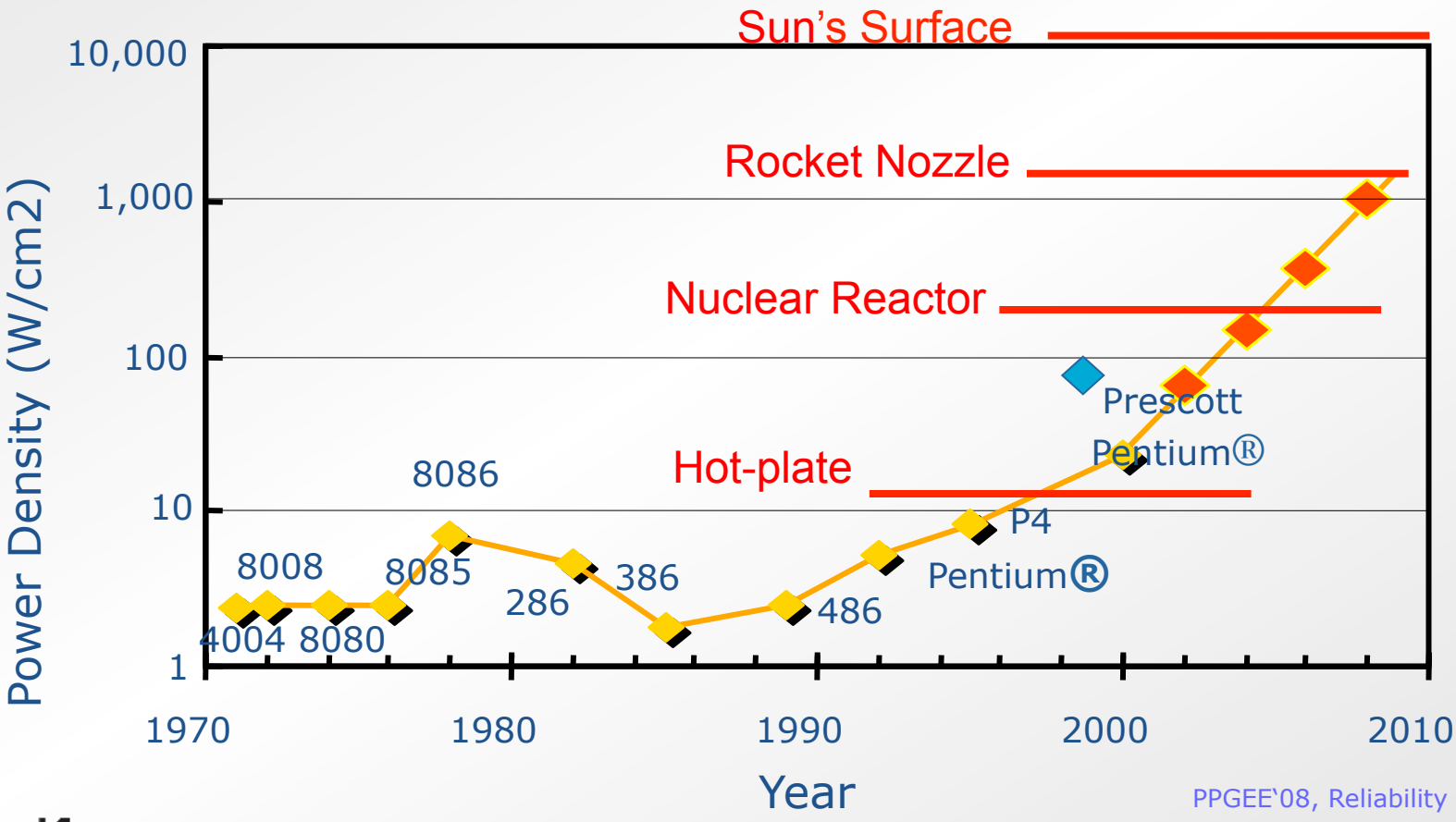
RESULTS:

Higher Density:	$\sim\alpha^2$
Higher Speed:	$\sim\alpha$
Power/ckt:	$\sim 1/\alpha^2$
Power Density:	~~Constant~~

- Approaching atomistic and quantum-mechanical boundaries
- **Atoms are not scalable!**

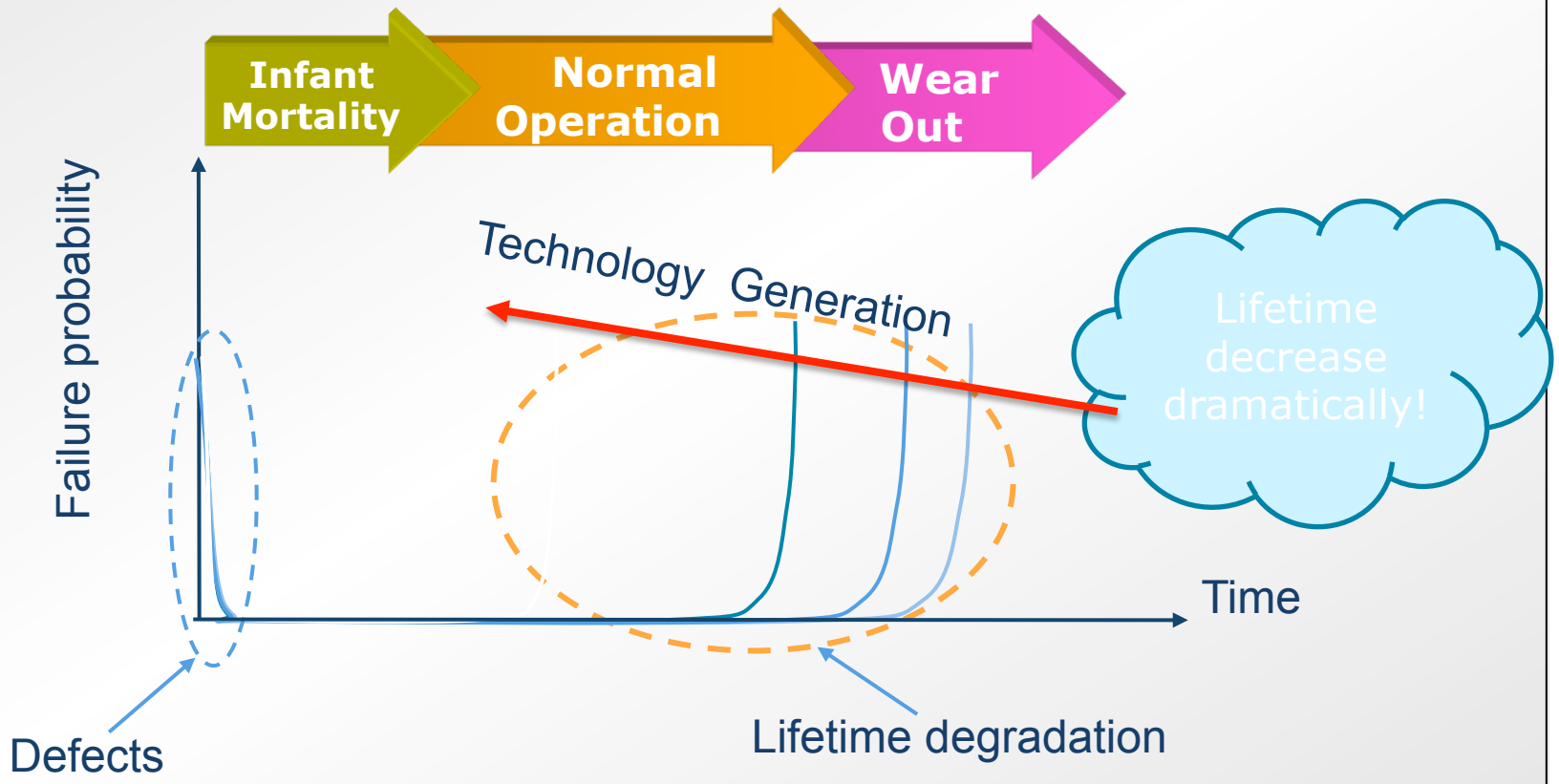
Power Density

Ever increasing power density leads to high in chip temperature, which accelerates temperature-elevated intrinsic failure mechanisms, e.g., Negative Bias Temperature Instability (NBTI).



Reliability vs. Technology

Device wear out is clearly observed and also End-of-Life approaches faster as technology advances.



"Bathtub" Curve of Reliability for Different Technology Generations

Reliability is Important For…

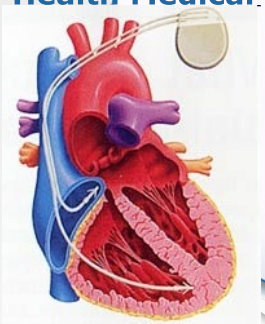
Air Flights

Military Service

Space Missions

Spaceship

Health Medical

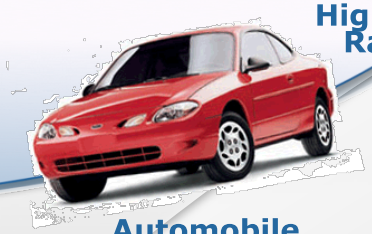


Almost every aspect of daily life

High Speed Railway

Enterprise Servers

Industry Control



Automobile

Consumer Electronics

TU Delft

Small Failures May Cost A Lot!

A small leak will sink a great ship.

http://telstarlogistics.typepad.com/telstarlogistics/2008/08/photos-and-vide.html

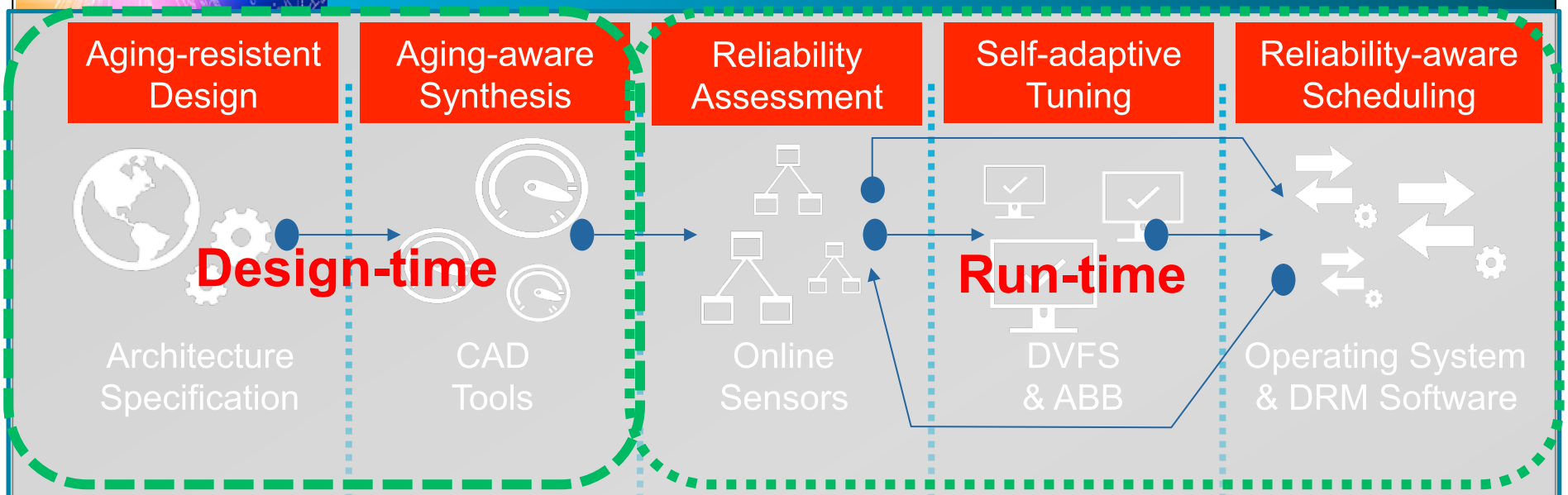
- B-2 bomber crash in Guam 2008
  - **$USD 1.4B loss**
- ***3 air data sensors malfunction***, moisture in the transducers during calibration distorted the information in the air data system.
- This caused the flight control computers to calculate inaccurate airspeed and negative angle of attack upon takeoff.

TU Delft

Contents

TU Delft

Reliability Aware Computing



Circuit Level Performance Degradation Model

Degradation Model of Device Parameter, like V_{th}, I_{DS} etc.

Fundamental Physics & Thermodynamics Model for Failure Mechanisms

NBTI | HCI | TDDB | EM | TC

Why DRM?

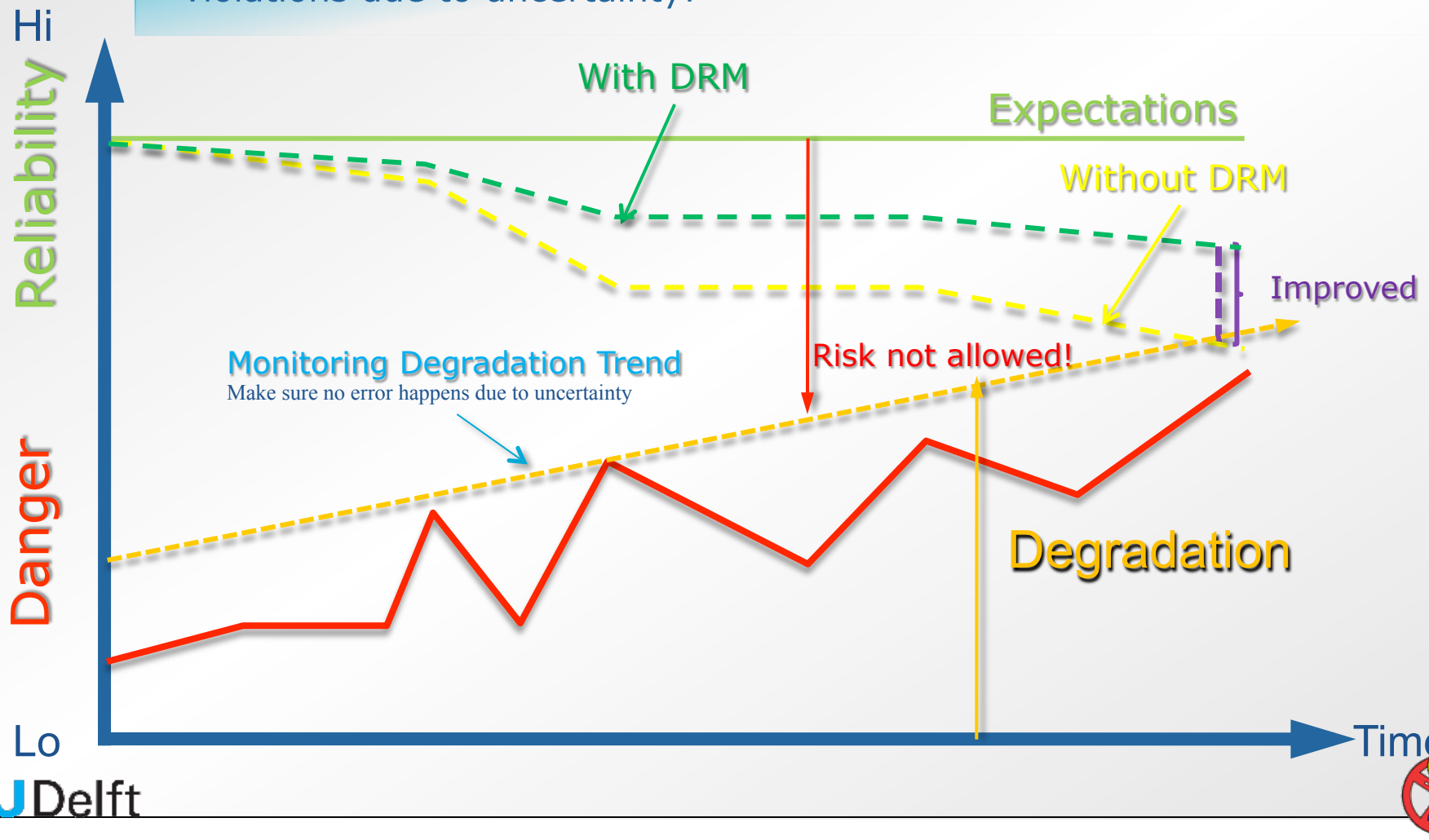
- **Expectations:** Desired work status defined at ***design-time***.
- **Practice:** Drifting Margin for fault-tolerance at ***run-time***.

* Adapted from Dekker, S. (2007), *The Field Guide to Understanding Human Error*.

TU Delft

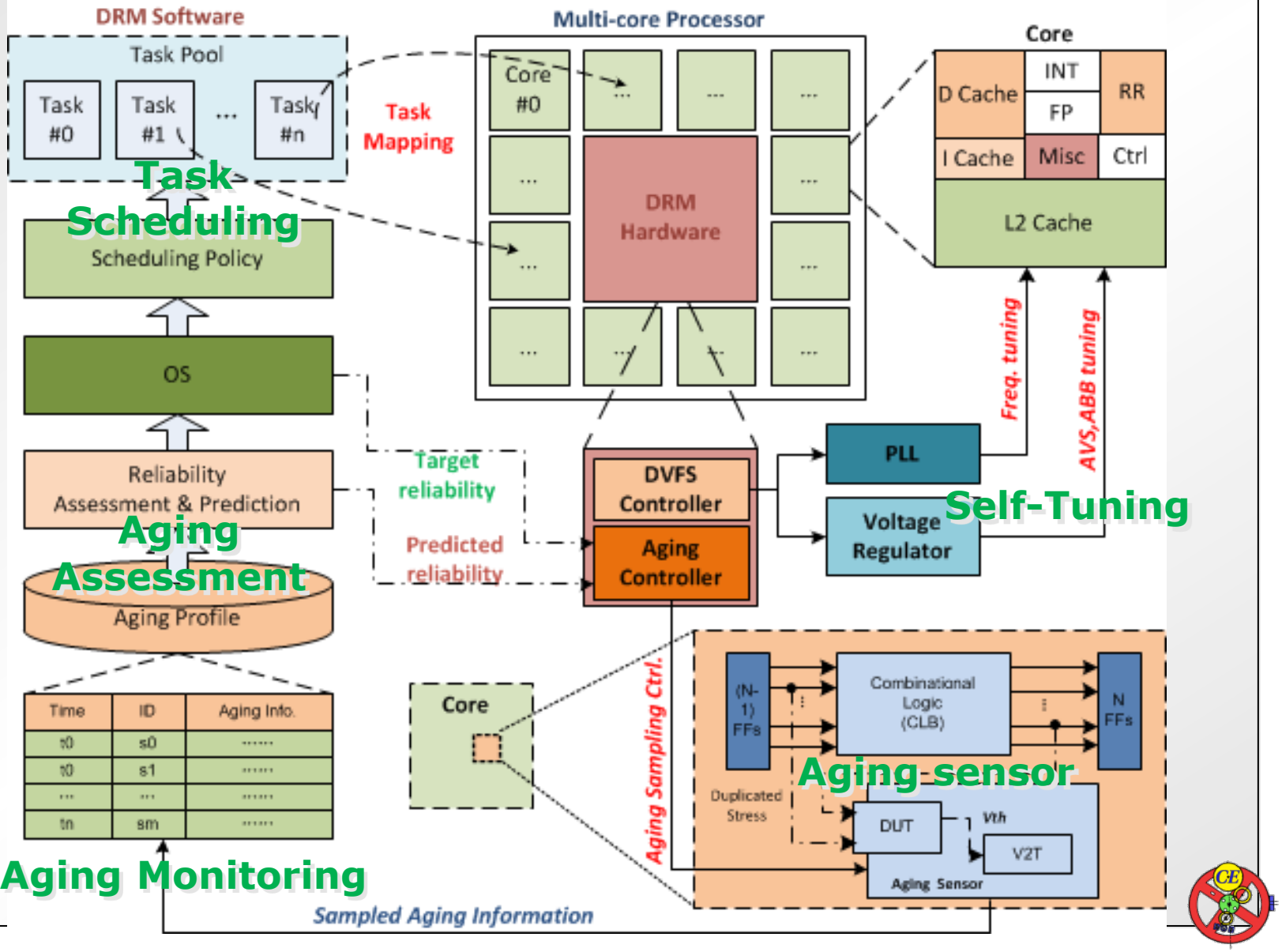
What DRM does?

DRM tries to slow down the reliability margin drift by all means, monitors the degrading trend of the system, and eliminates risk and violations due to uncertainty!



TU Delft

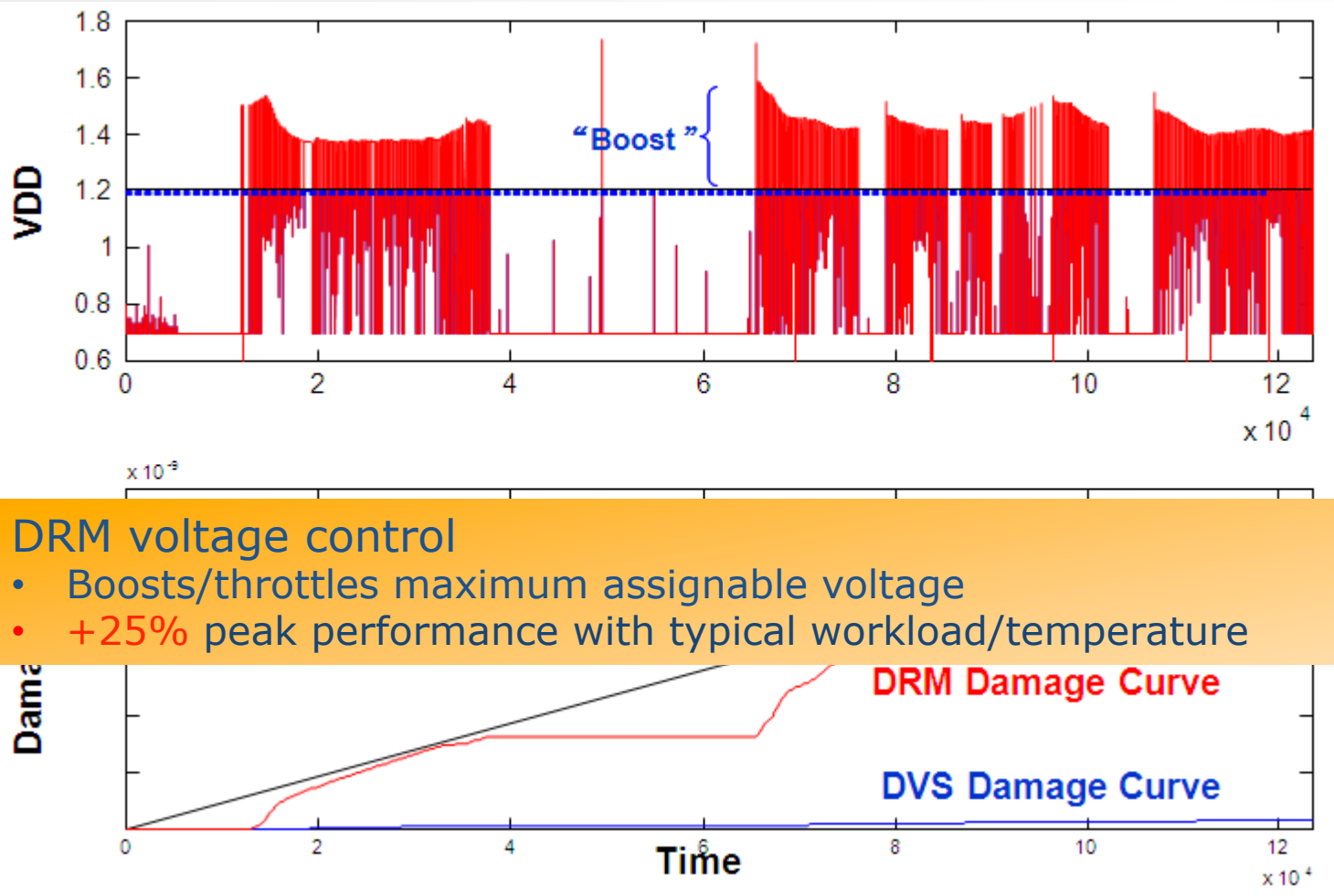
A DRM Framework



TU Delft

Impact of Dynamic Reliability Management

In a DRM System, the maximum voltage can be "*boosted*" to allow periods of higher peak performance while maintaining a margin below the budgeted damage curve.



DRM voltage control

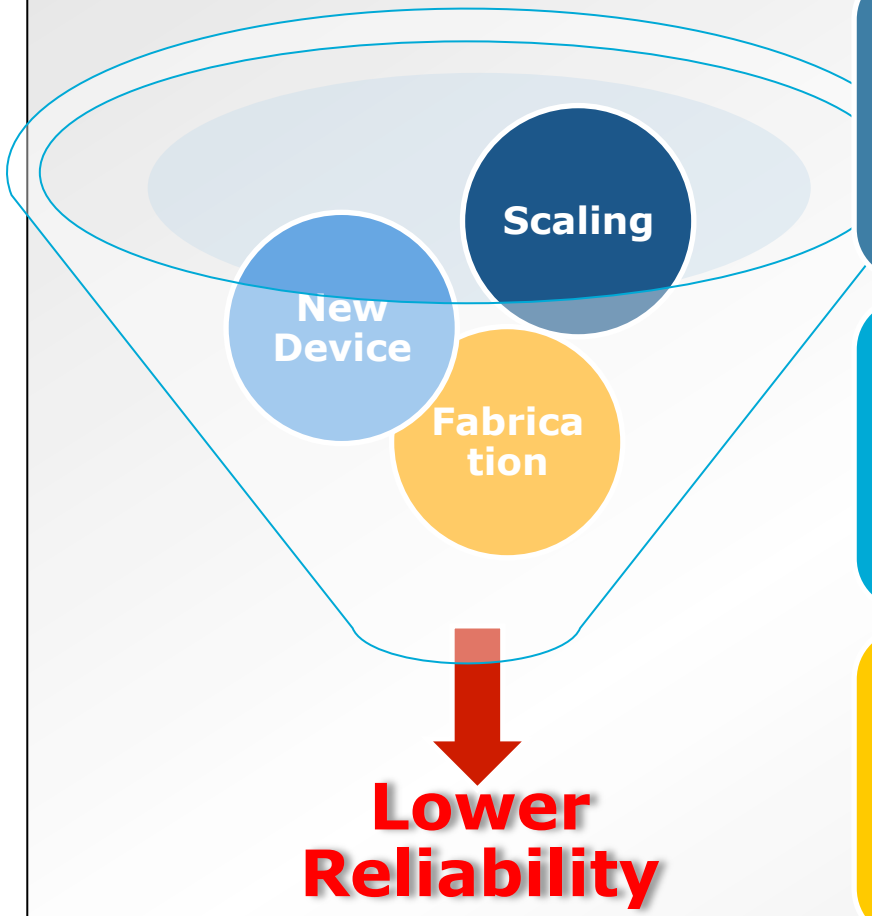
- Boosts/throttles maximum assignable voltage
- +25% peak performance with typical workload/temperature

TU Delft

Contents

TU Delft

Reliability Becomes More Difficult



Scaling

- Device dimension decrease continuously, physical limit is approaching.

New Device

- New device structures and materials introduces new reliability concerns.

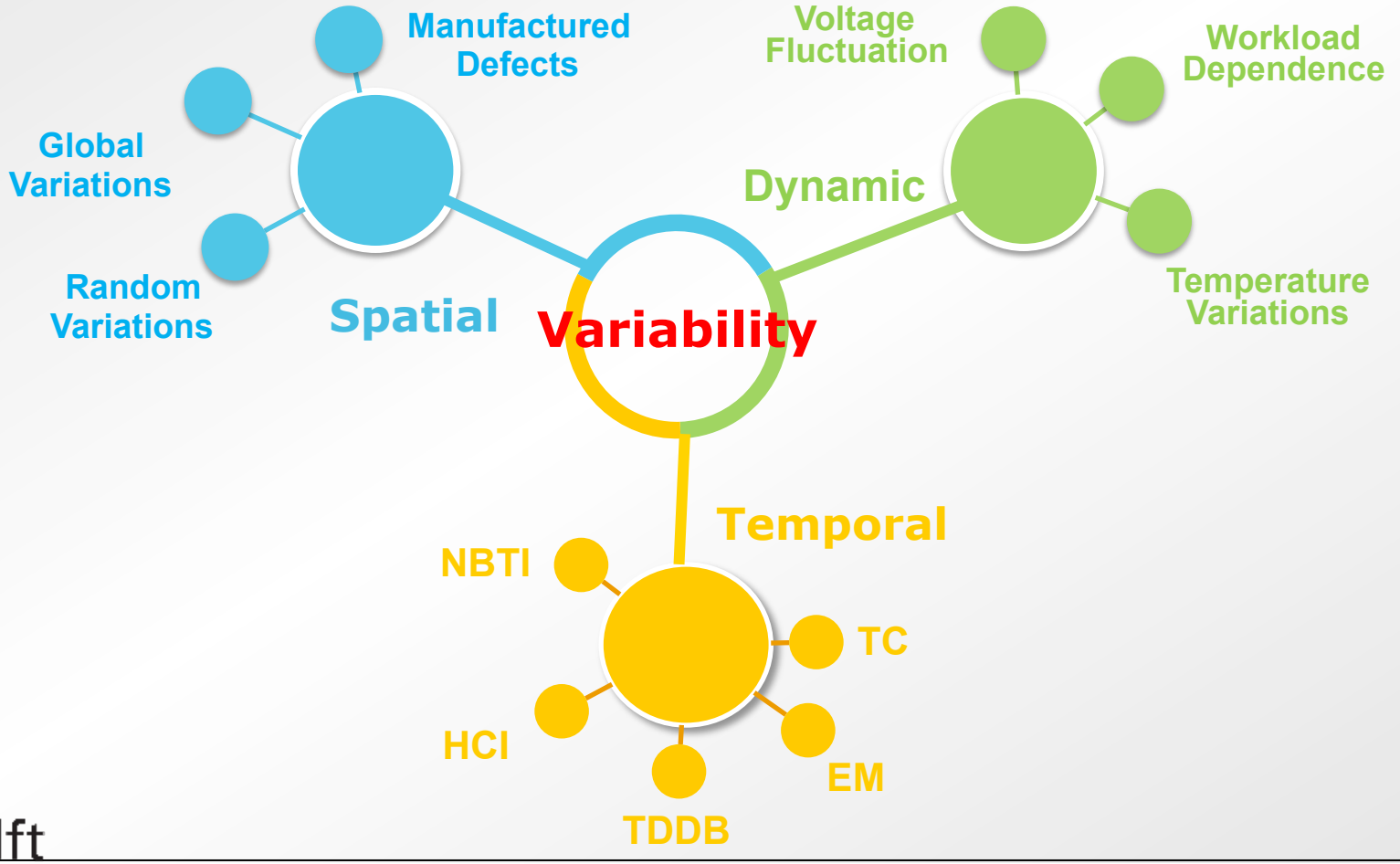
Fabrication

- Manufacturing variability increases since process geometries gets smaller, optical diameter doesn't decrease proportionally.

TU Delft

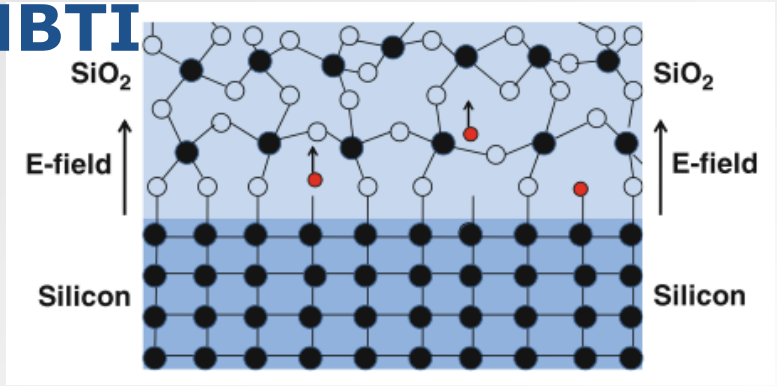
Reliability Challenges

The most challenges of reliability comes from **Variability**, which can be further divided into three categories: ***spatial***, ***temporal*** and ***dynamic*** variability.



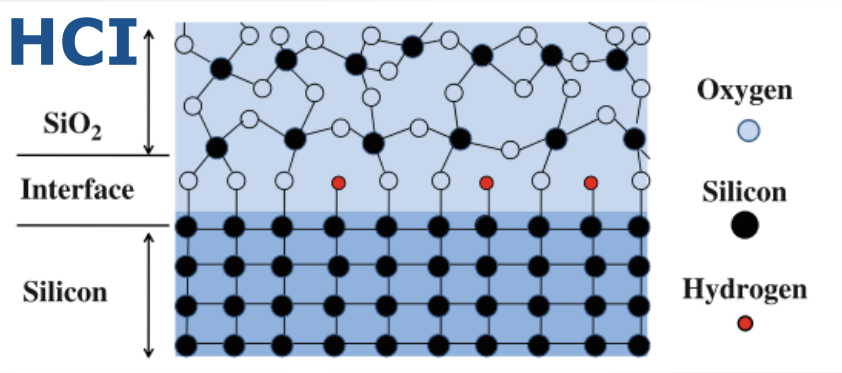
Failure Mechanisms

NBTI



- Holes from the inversion layer tunnel into the gate oxide, break the Si-H bonds and leave behind interface traps.
- H atom diffuses away from the Si/SiO2 interface.

HCI



- "Lucky electrons'" gain enough energy while drifting across the channel.
- The "hot" electrons produces interface damage in a localized region near the drain end.

- NBTI and HCI creates hole traps at Si/SiO2 interface and in the oxide, which leads a *positive shift* of the device $|V_{th}|$;
- NBTI and HCI highly depends on the stress probability at the device;
- Furthermore, NBTI and HCI are prone to voltage fluctuation and temperature variations.

TU Delft

V_{th} Temporal Degradation

- **V_{th} Degradation under NBTI**

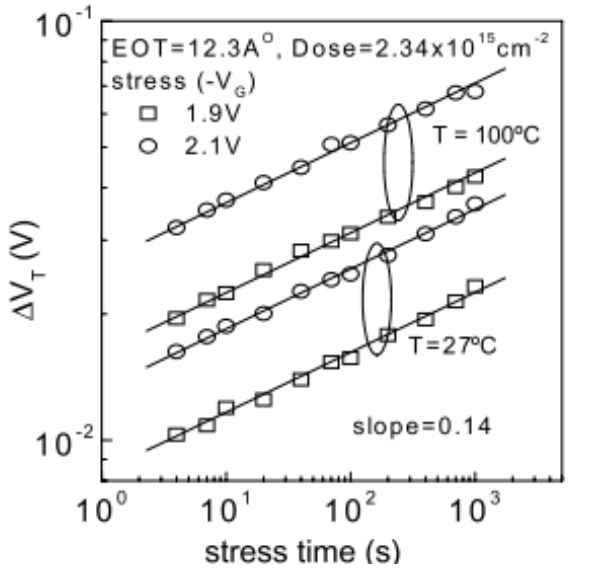
DC stress

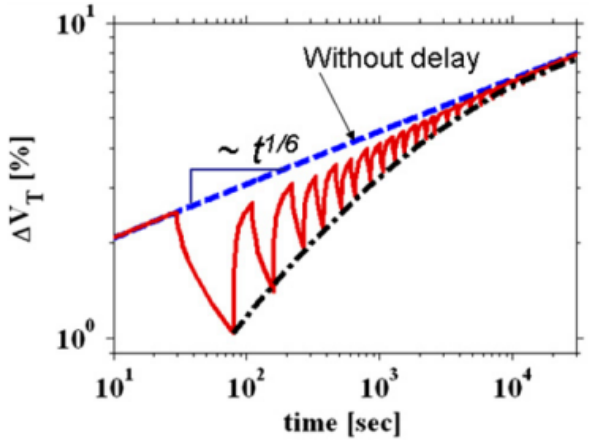
$$\Delta V_{th}(t) = (1+m)\frac{q\Delta N_i t(t)}{C_{ox}}$$

AC stress

stress $\Delta V_{th} = (K_v(t-t_0)^{1/2} + \sqrt[2n]{V_{th0}})^{2n}$

recovery $\Delta V_{th} = V_{th0}(1 - \frac{2\xi_1 t_e + \sqrt{\xi_2 C(t-t_0)}}{2t_{ox} + \sqrt{C}t})$





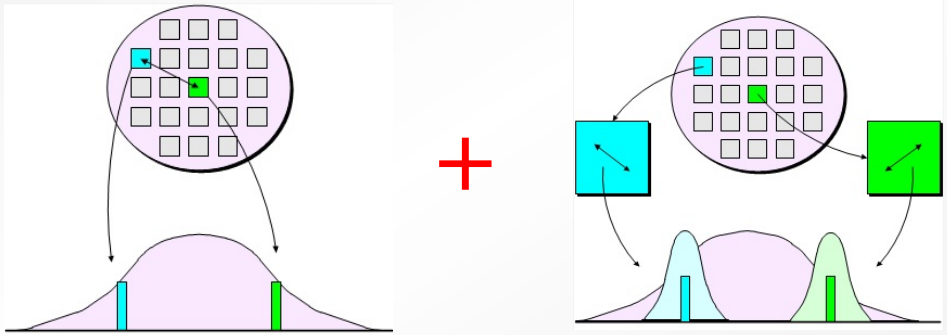
Alam et al. MR. 2006

TU Delft

Variability Correlation

- ***Variation sources of V_{th}***

$$\sigma_{V_{th},PV} = \frac{A_{VT}}{\sqrt{WL}}$$

Process Variation =  +

And

Temporal Variation
(by NBTI)

$$\sigma_{NBTI(t)} = \sqrt{\sigma^2_{\Delta V_{th}(t)} + \sigma^2_{\Delta V_{th}(w)}}$$

= **Spatial** + **Dynamic**

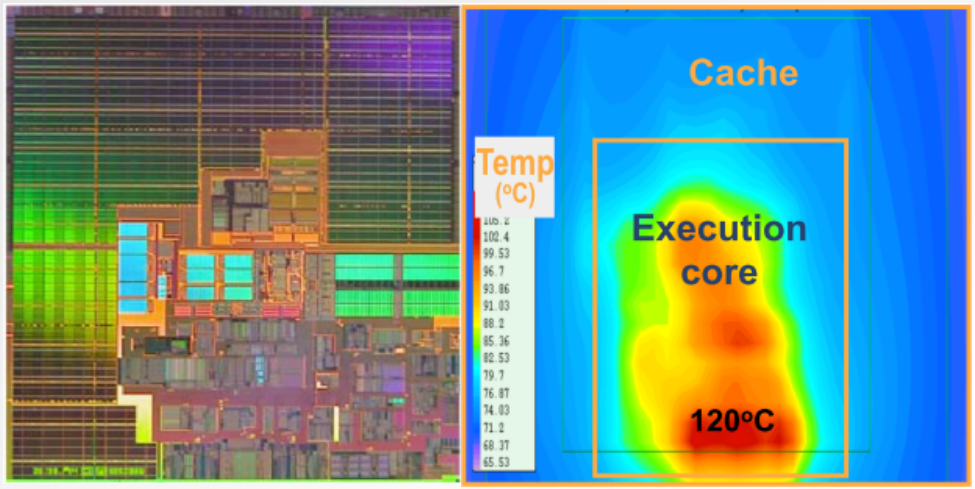
Random dopant ***Workload***

Correlations between different variability sources make it complicated to build accurate physical models for failure mechanisms!

TU Delft

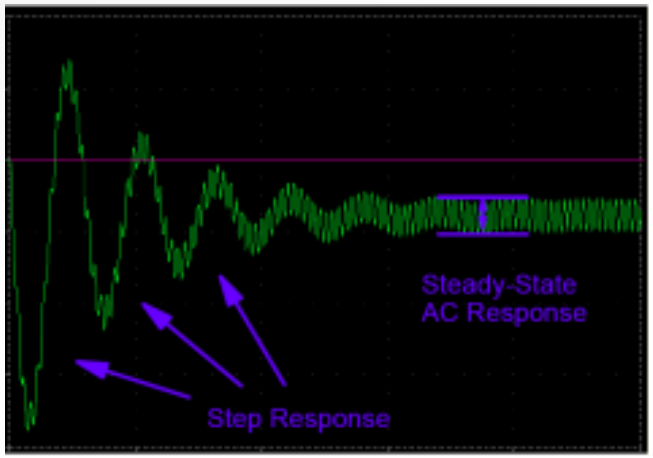
Dynamic Variations

Thermal Map (Itanium@1.5GHz)



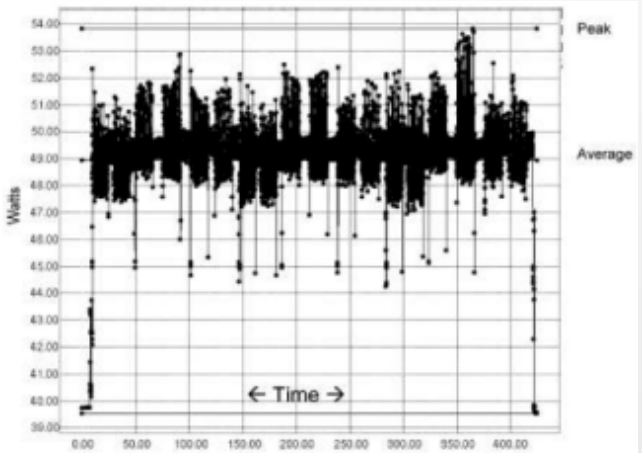
Dynamic variations caused by workload are hard to predict!

Voltage Variation



Source: D. Hathaway, SLIP 2005

Power Variations



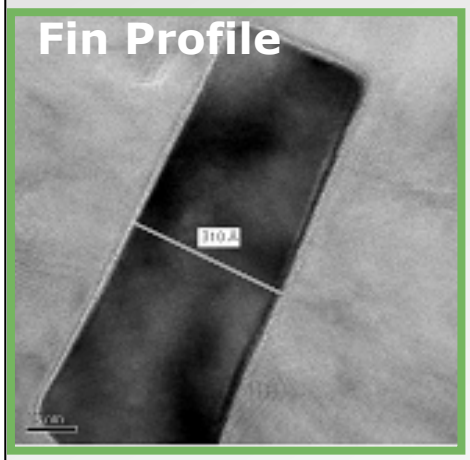
Source: Naffziger *et al*, JSSC 2006

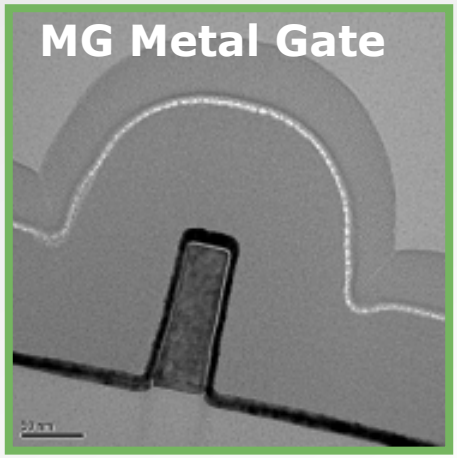
TU Delft

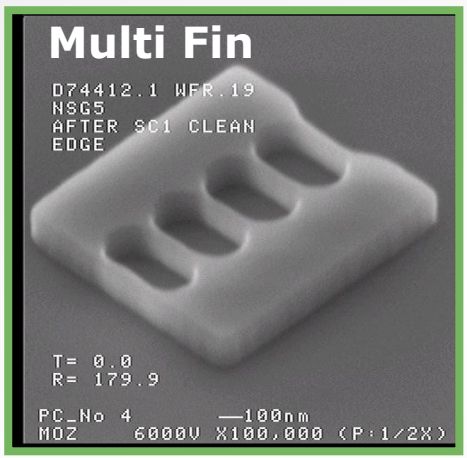
Contents

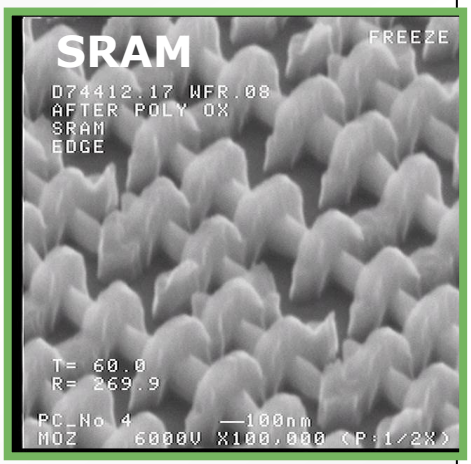
TU Delft

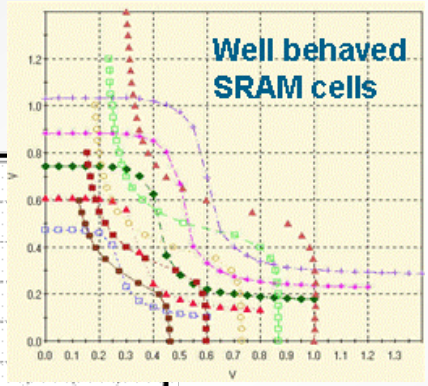
New Devices

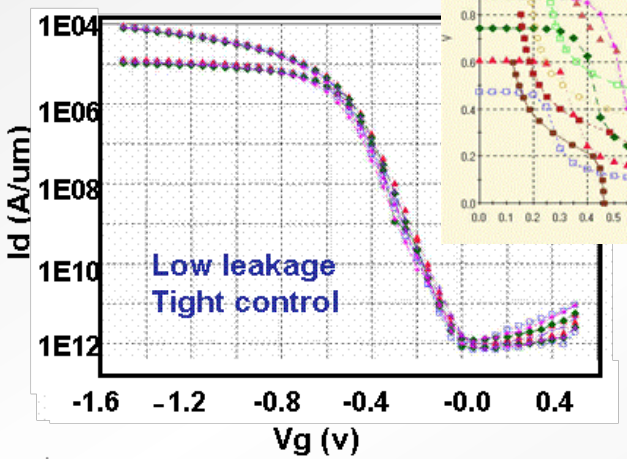








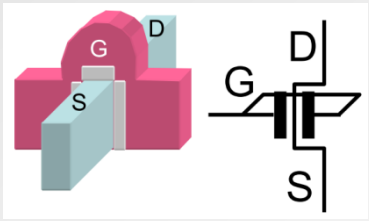




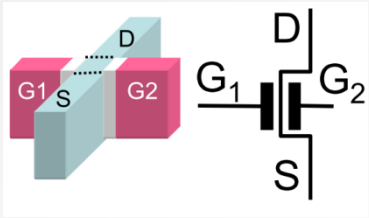
- **FinFET integration showing:**
  - Single Gate
  - Multi Gates
  - Independent Gates
  - Midgap Metal Gate and undoped channel (FD) Devices
- **Calibrated circuit simulation show 34% F_{max} increase at 1.1V**
- **Enable new circuits and application beyond the planar devices**

TU Delft

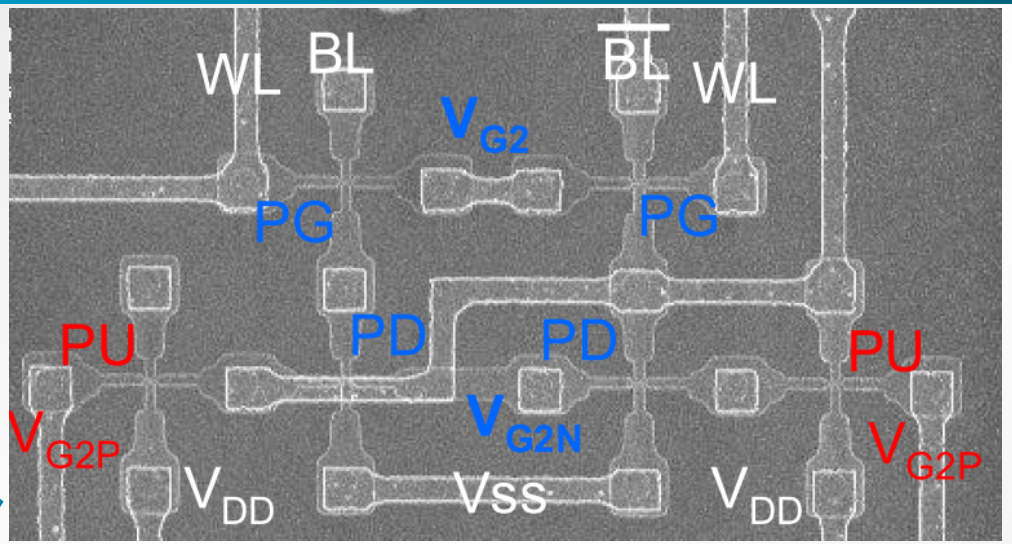
IG-FinFET SRAM Cell



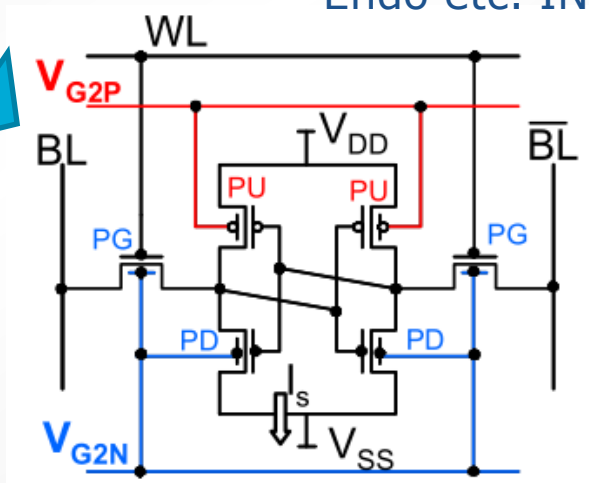
Common-Gate FinFET



Independent-Gate FinFET



Endo etc. INEC. 2011



NBTI mitigation is possible by using IG-FinFET!

IG-FinFET 6T SRAM Cell:

(a) allows to control V_{th} for SRAM cells;

(b) no significant area & power consumption trade-off introduced.

As Scaling Continues…

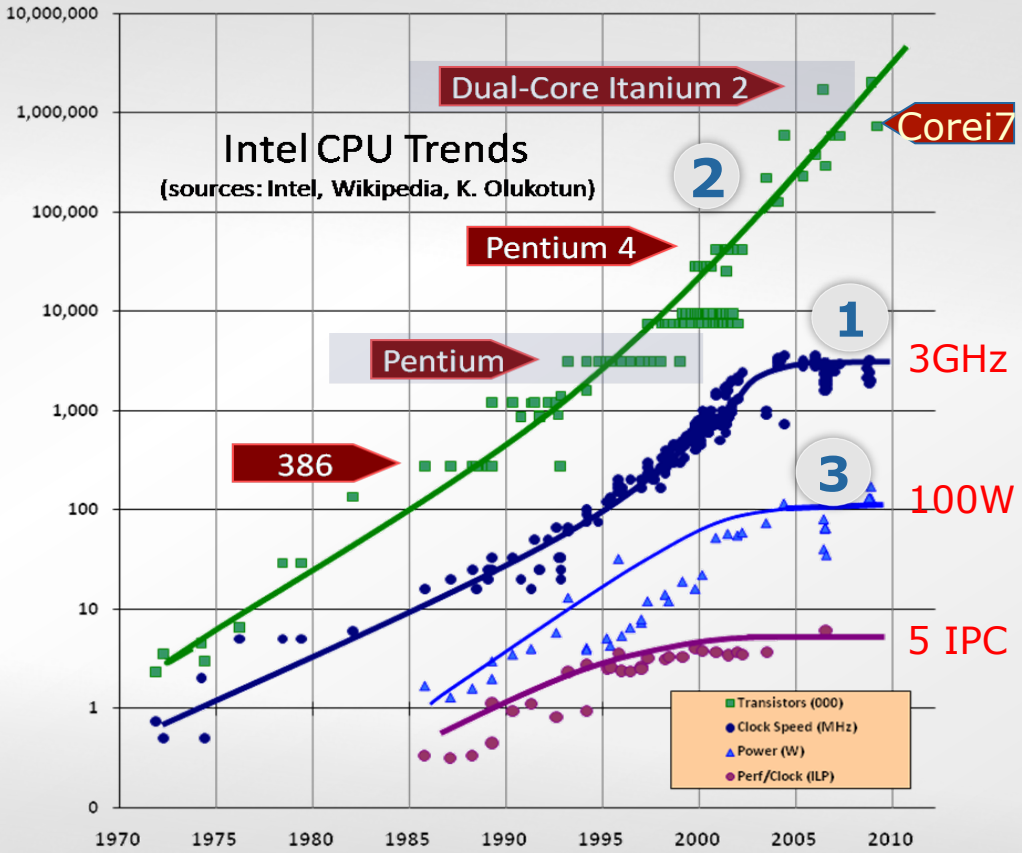
Multi/Many core processor is more power efficient, and give more flexibility to perform DRM with graceful degrading policy!

1. **Clock Frequency**
   Fixed or getting slower.

2. **# of Transistors per Chip**
   Doubles every 18-24 months.

3. **Power per Chip**
   Reaches almost maximum.

Moore's Law
Reinterpret to *# of transistors or cores per chip*.



From K. Olukotun, L. Hammond, H. Sutter, and B. Smith

TU Delft

Summary

- Reliability is getting more important
  - Feature size approaches physical limits thus devices are unreliable than ever due to their small size;
  - Power density leads to high temperature and electric field on chip, which accelerates the aging progress.
- Reliability management is more complex/difficult
  - Three types of variability exists, i.e., spatial variations, temporal variations (wearout), and dynamic variations, which create many reliability uncertainties;
  - Different types of variability are correlated, thus reliability models are getting more complicated.
- Dynamic Reliability Management (DRM)
  - Slow down the degradation progress by performing, e.g., reliability-aware resource allocation;
  - Can boost performance within a certain reliability margin;
  - Can provide End of Life prediction & alarms;
  - Key components
    - Physics Models, Aging Sensors, Reliability Assessment;
  - Strategies
    - Self-tuning (DVFS, ABB, etc.), Task Scheduling, …