Dependability Challenges at the Device and Circuit Levels

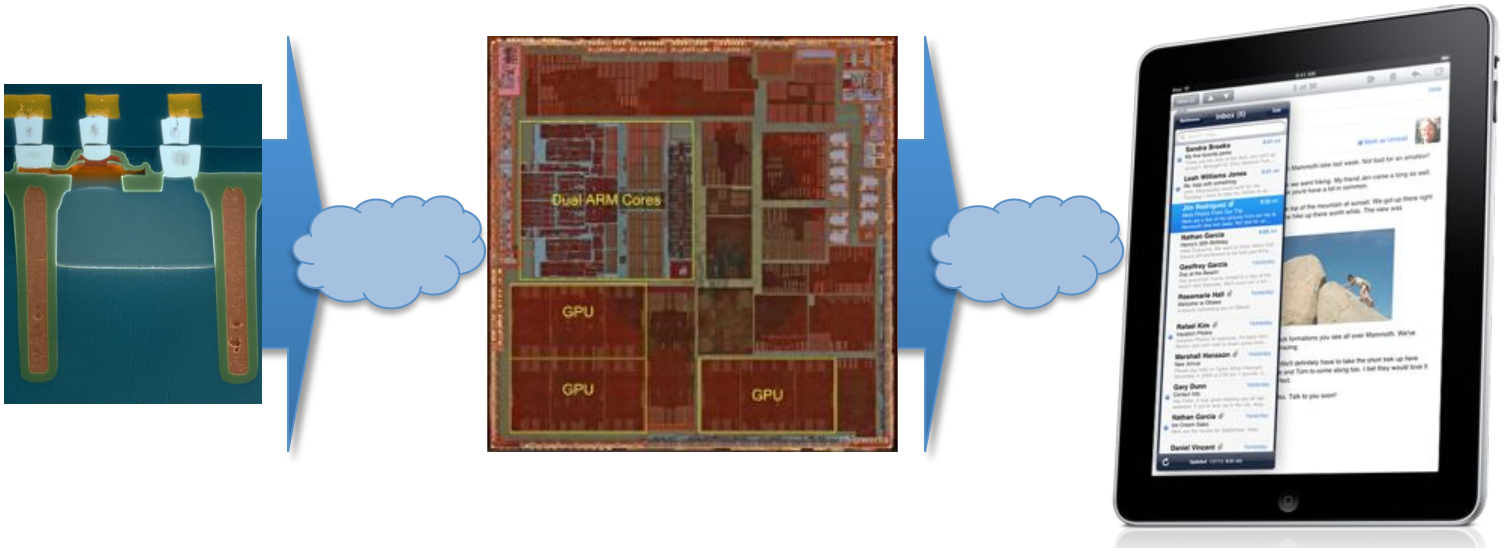
Sani R. Nassif

IBM Research – Austin
December 1st, 2012

Abstraction: A Cornerstone of VLSI

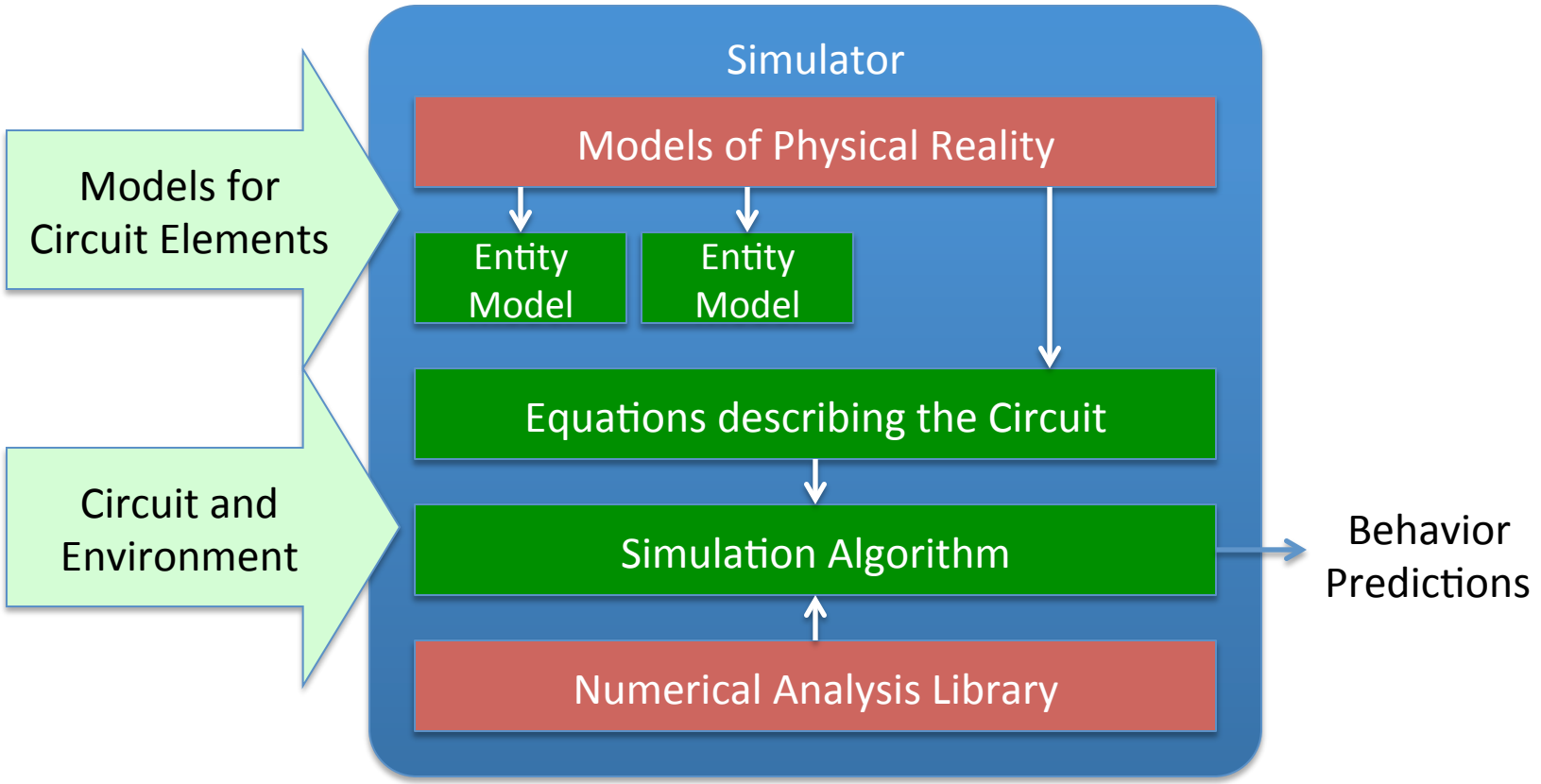
- Modern integrated circuits have Billions of components.
  - A Boeing 747-400 has “just” 6 Million parts!

- VLSI engineering would not exist if not for the ability to perform abstraction to hide detail.
  - Many levels of abstraction between “device” and “application”.



Circuit and Device Abstractions for Simulation

- The primary purpose for abstraction is **prediction**.
  - VLSI engineers rely exclusively on simulation (prototyping rarely done).
- Simulators take mathematical models of physical reality and use numerical analysis to solve those models in order to make predictions.

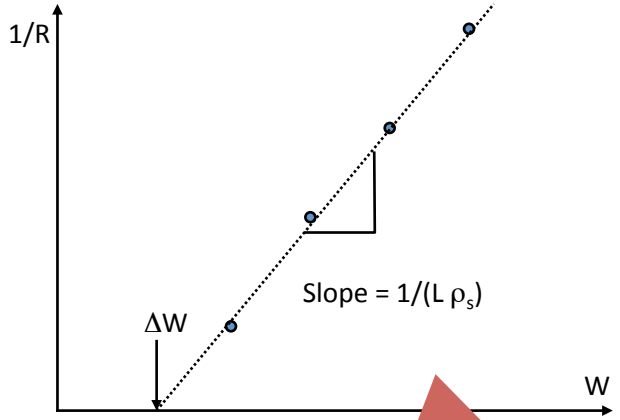


Device Abstraction: Simulation Models

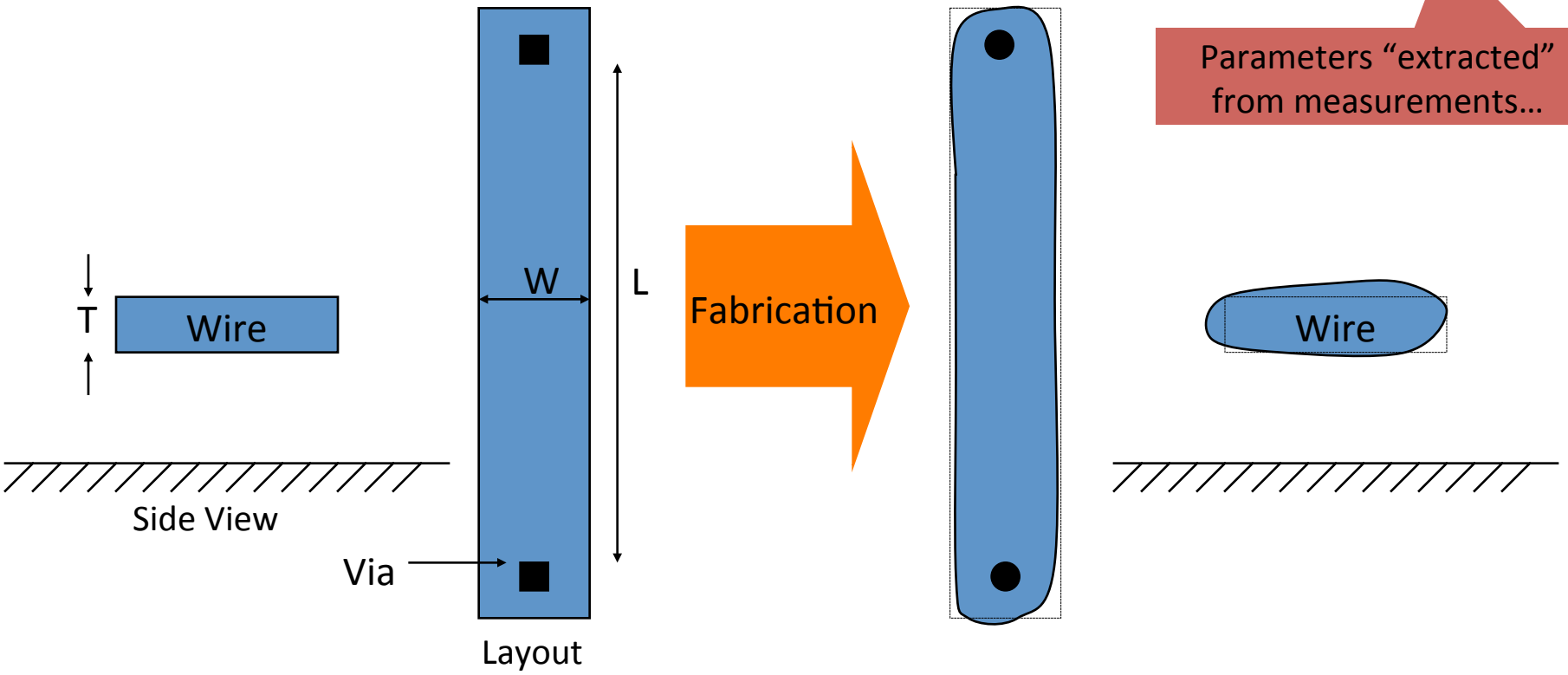
Simplest case… wire resistance:

$$R = (\rho / T)\,(L - \Delta L) / (W - \Delta W)$$

ρ/T = "sheet resistance"

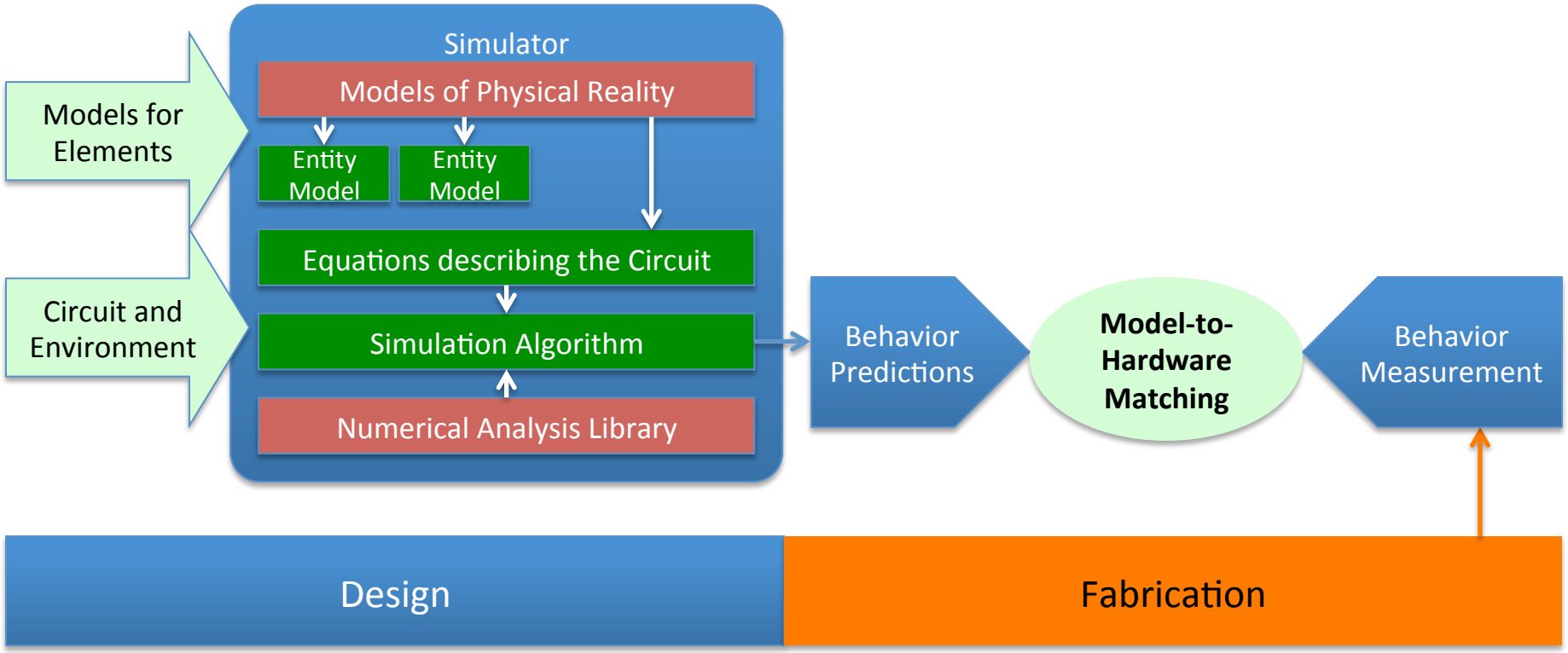


Parameters "extracted" from measurements…



Reminder: Simulation is for Prediction!

- We use simulators, and the models of devices within them, in order to predict the behavior of circuits before they are fabricated.
- We make sure the models are correct by corroborating with respect to measured circuits.



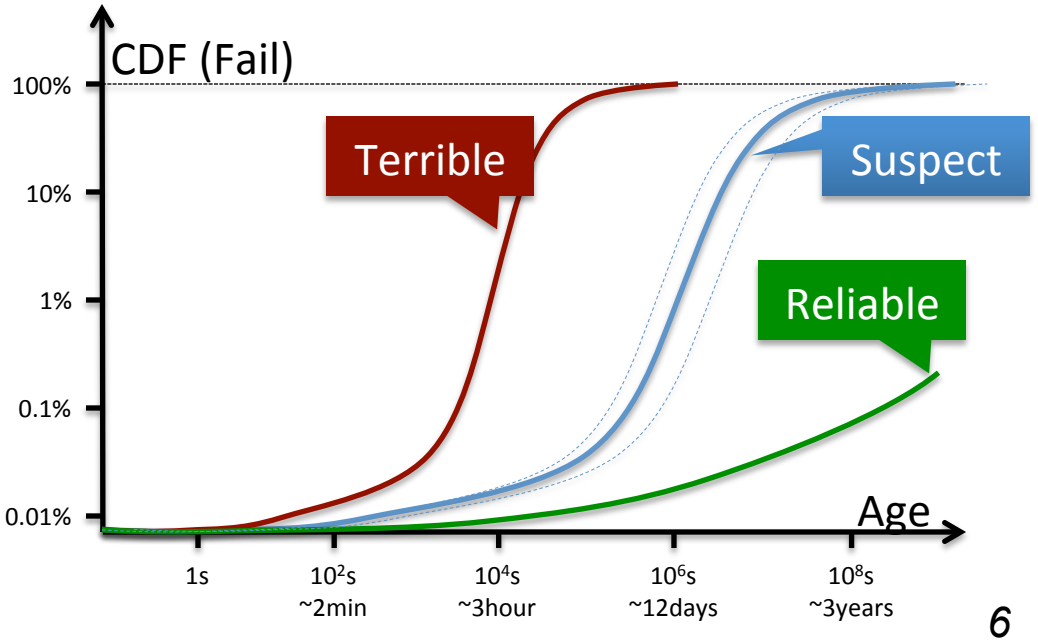
Dependability = Predictability

- Designers value *predictability* over all else…
  - Lack of predictability ➔ Lack of dependability!

- Phenomena which can be predicted can be margined against and eventually somehow can be “fixed”.
  - Unless the impact is so significant that no fix is possible.

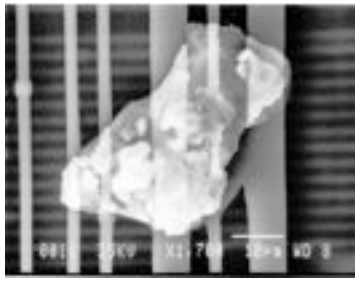
- Example: wire lifetime.

- Key idea: uncertainty mixes badly with lack of dependability.



Traditional Dependability Models

- Until relatively recently (~130nm) the distinction between hard and soft failures was relatively clear.
  1. Catastrophic/Hard faults were dominated by defects.
  2. Parametric/Soft faults were dominated variability from the manufacturing process.



- Typical action plan: clean up the fab and your functional yield goes up. Then tighten the tolerances and your parametric yield goes up.

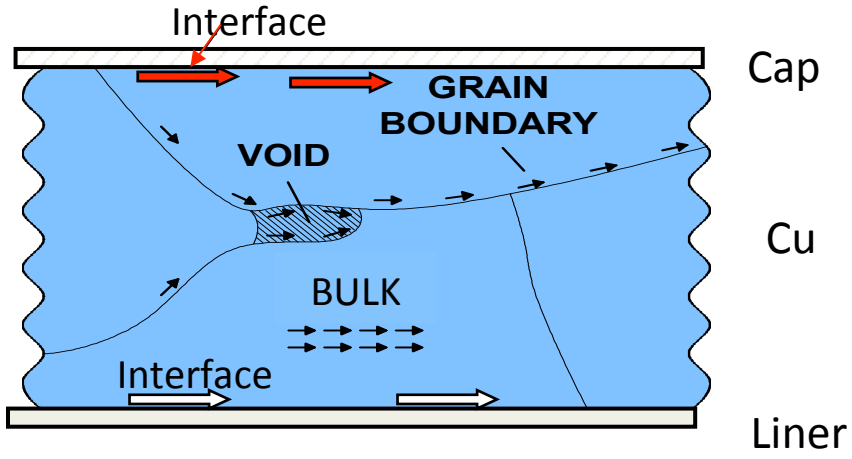
- This is no longer the case… numerous systematic phenomena can cause uncertainty and failure, and therefore dependability challenges.

Linking Uncertainty and Dependability

- Consider again the case of wire lifetime.
  - Can estimate via 1st principles MC carrier transport.
  - Mean Time to Failure semi-empirical model: $\mathrm{MTF}(t_{50}) = \frac{A}{j^n} e^{\frac{H}{kT}}$

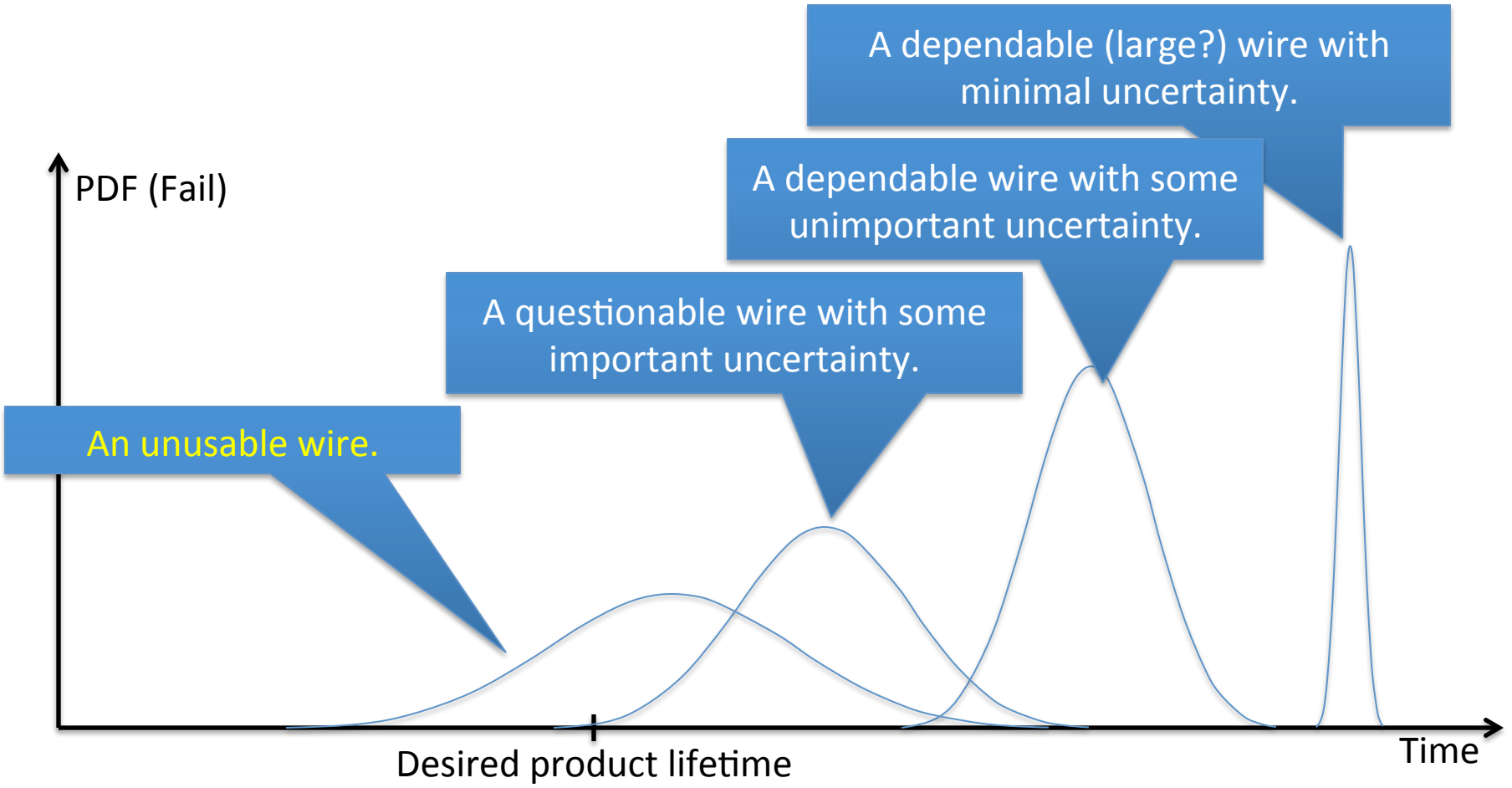
- Sources of uncertainty:
  - **Operating conditions**: current and temperature.
  - **Microscopic structure**: grain shape and boundaries.

- In modern technologies, the impact of structure is mounting as dimensions approach the mean-free-path of ē in Cu.
  - Mean free path ~ 40nm.
  - Implication: increased lifetime uncertainty for wires!



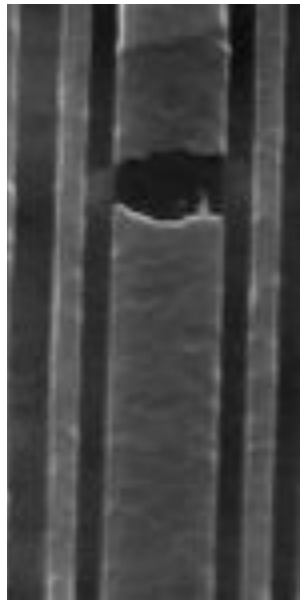
Wire Lifetime / Dependability

- Because of intrinsic variability, failure time is random.
  - MTF is “mean”… There is also “STF” for standard deviations!



Challenges in Dependability for Circuits/Devices

- The major challenge is to allow **prediction of failure**.
  - This will require appropriate abstractions at the device, circuit, gate, macro, unit, chip, software, system, etc…

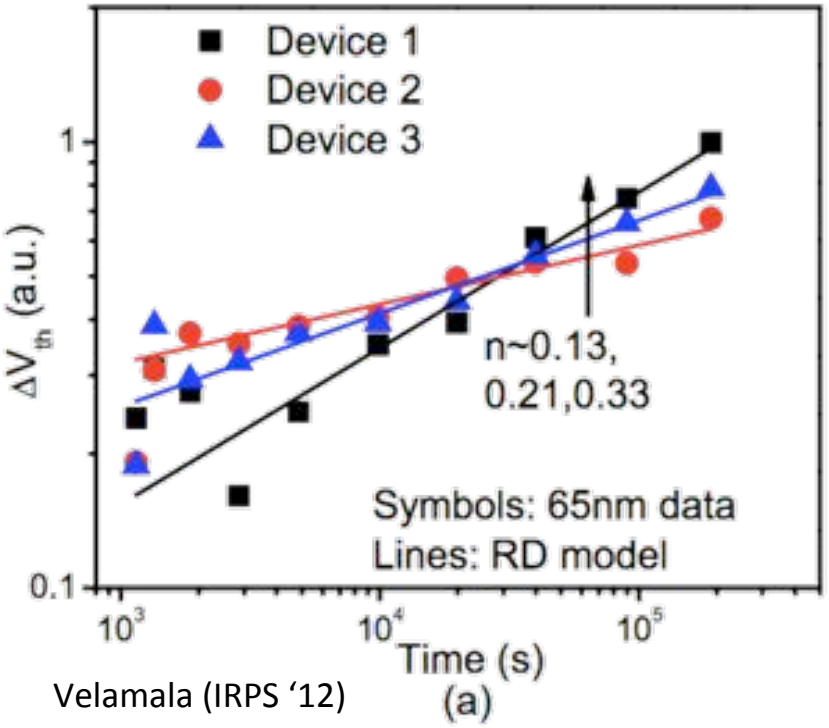


- A second challenge is to **incorporate uncertainty**.
  - When failure is improbable, uncertainty is not difficult. Crude estimates are OK.
  - When failure is more probable, it becomes critical to bound the uncertainty!

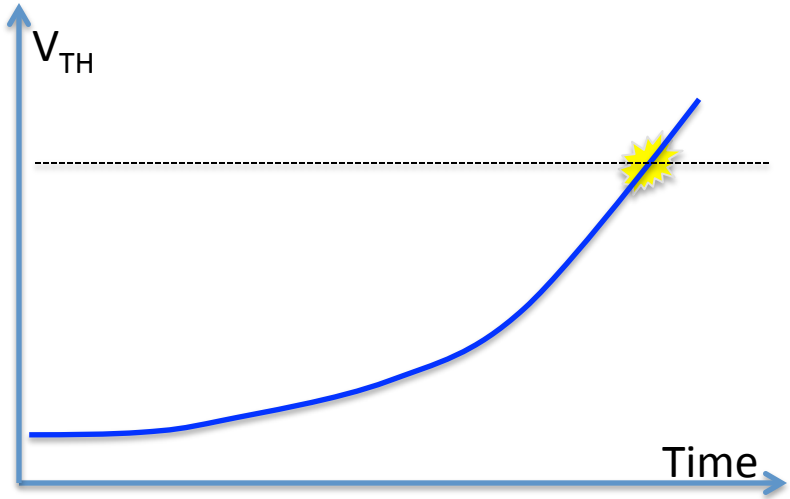
- The circuit/device level must present dependability limits to higher levels of abstraction to enable cross-layer solutions.

Dependability at the Device Level

- For usable devices, time-zero failure is not an issue.
  - We must assume that manufacturing has done its job…
- Devices "fail" due to a number of known stress-related phenomena. Failure may mean "going out of spec".
  - Example: Negative Bias Temperature Instability.

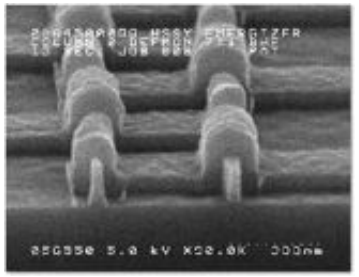


Velamala (IRPS '12)



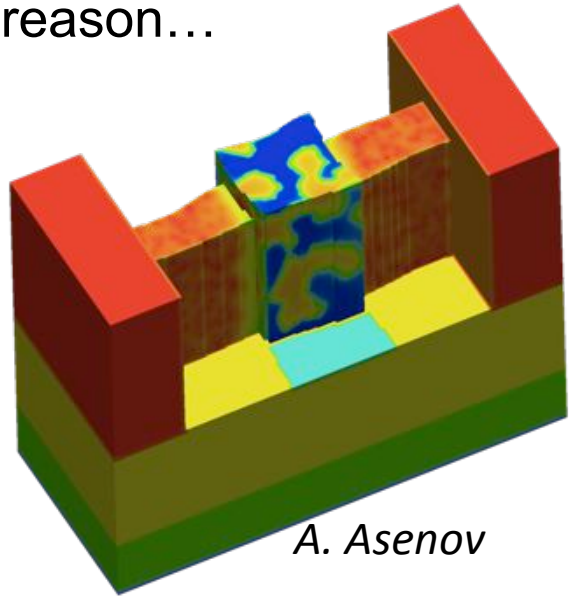
Dependability at the Device Level

- Accurate Spice models are an **absolute necessity**.
  - Understanding the impact of manufacturing variability is a plus.
  - Example: interaction between NBTI and VT variations.



- But… operating environment is also important!
  - Lack of knowledge ➔ excessive pessimism.
  - We can no longer afford to add margin without reason…

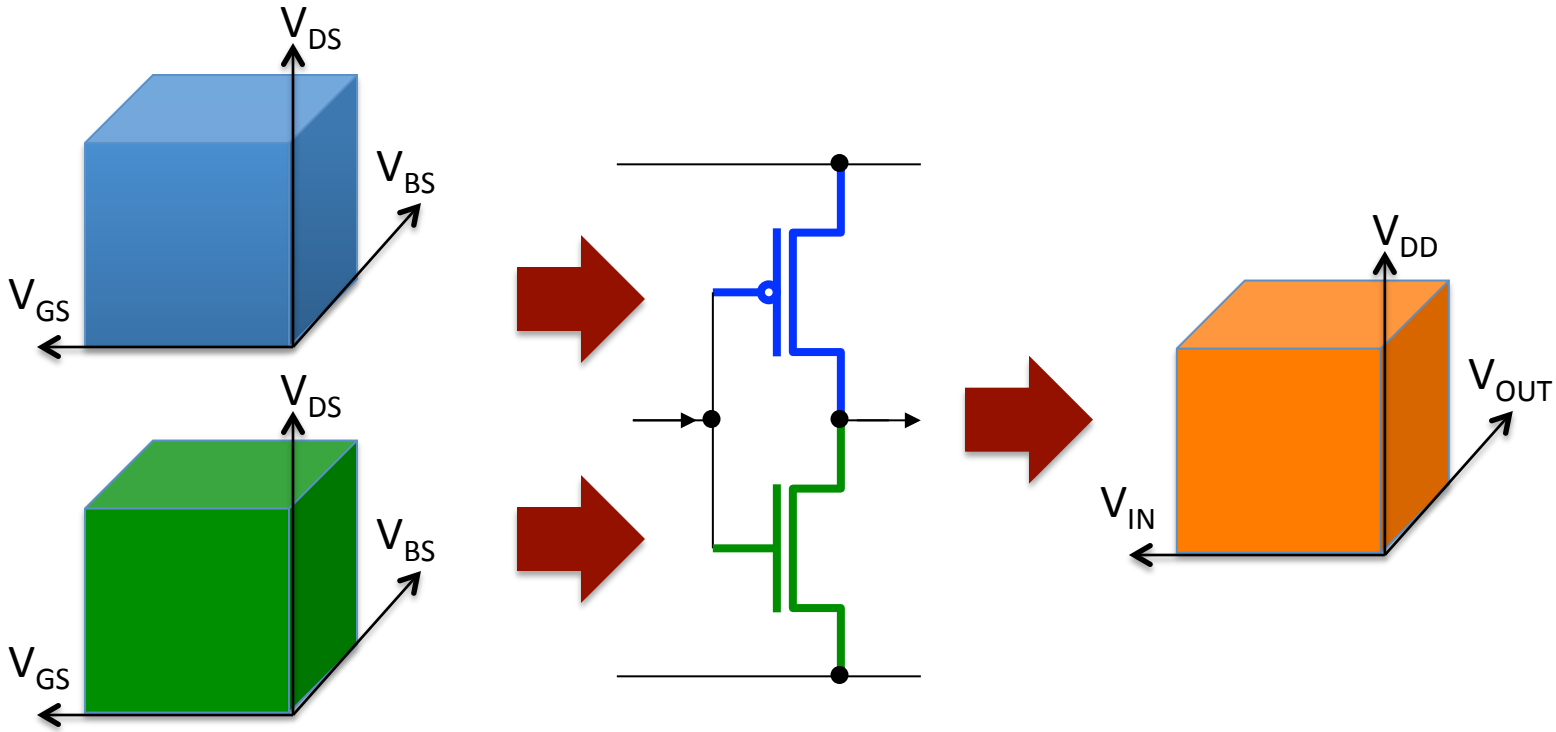
- Upcoming challenge: understanding device degradation for new device types, like FinFET.
  - Models relatively simple and inaccurate.
  - Insufficient history.



A. Asenov

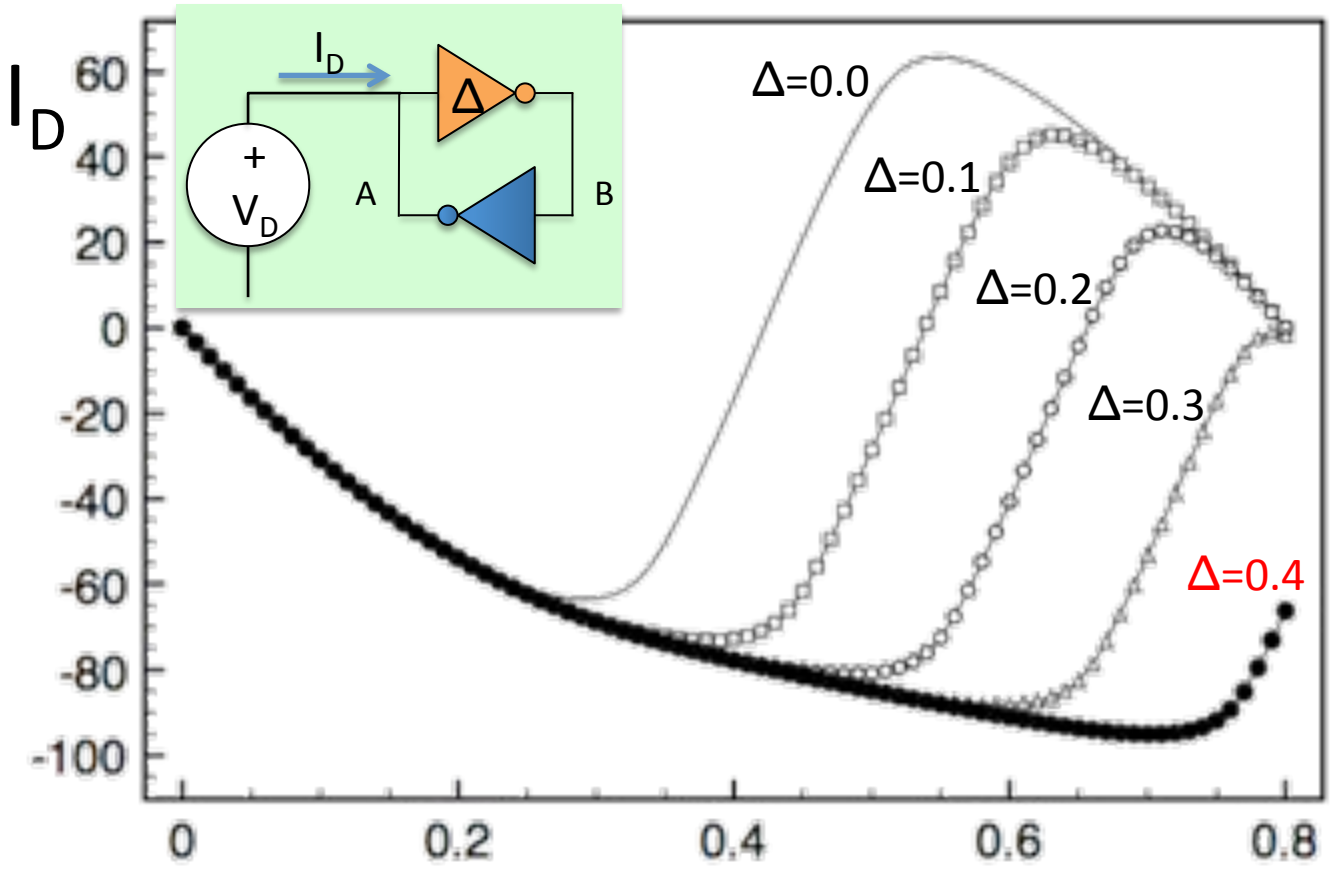
From Devices to Circuits

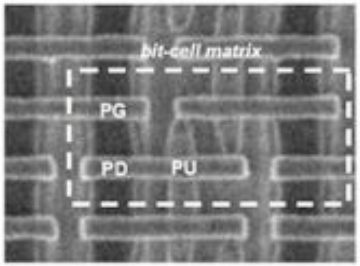
- A good robust design methodology would predict circuit behavior from device behavior (via simulation).
  - Circuits cause a reduction in the operating range for devices.
- But… Circuits need to be verified over the <u>complete</u> parametric space of constituent devices!

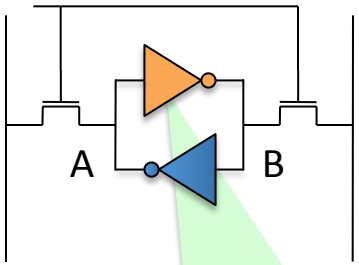


Dependability at the Circuit Level

- When does a circuit fail with all devices not failing?
  - Most probably when designer does not verify its operation over the complete parameter range (process/aging window)!







Orange inverter VT will shift due to NBTI
(maybe A = 0 for too long)

Understanding Circuit-Level Dependability

Two major components:

- Good device models.
- Robust design methodology.

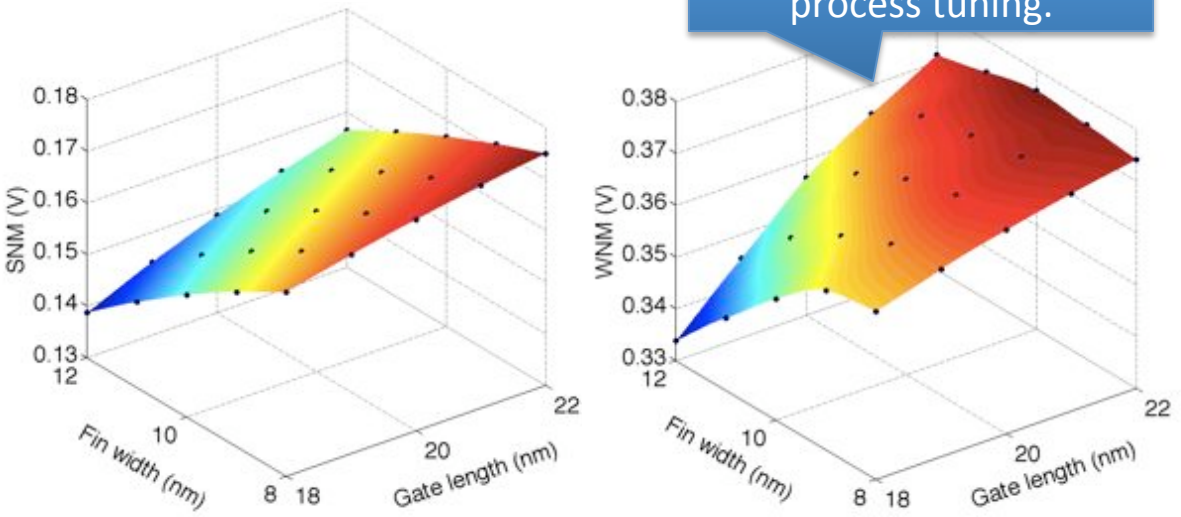
- This is a continuing area of challenge.
  - Accurate models are usually “late”.
  - Reliability models are usually “very late”.
  - Little support from EDA companies.

- One method to accelearate: “TCAD-based Path Finding”.
  - Use Technology CAD to make early predictions and determine problem areas.

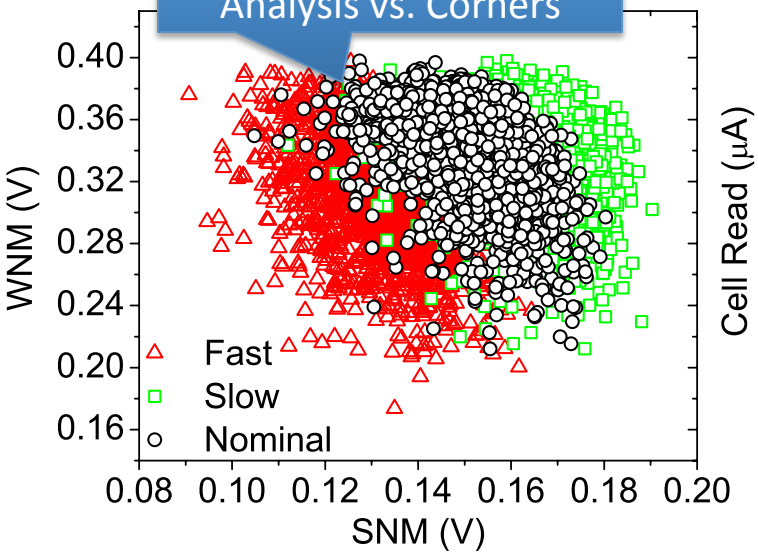
FinFET SRAM Variability Simulation

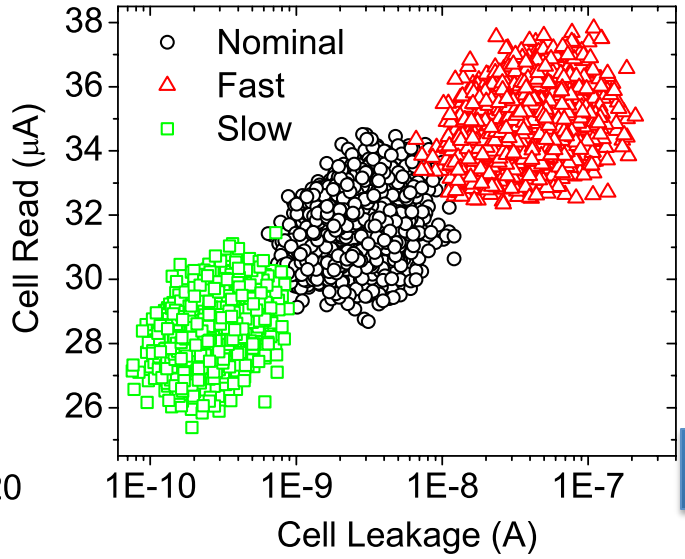
- Joint project with U. Glasgow on 14nm FinFET "path finding" to enable rapid yield ramping.

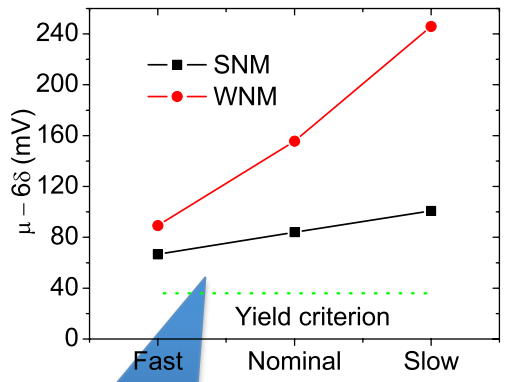
Nominal analysis vs. process tuning.



Statistical Performance Analysis vs. Corners







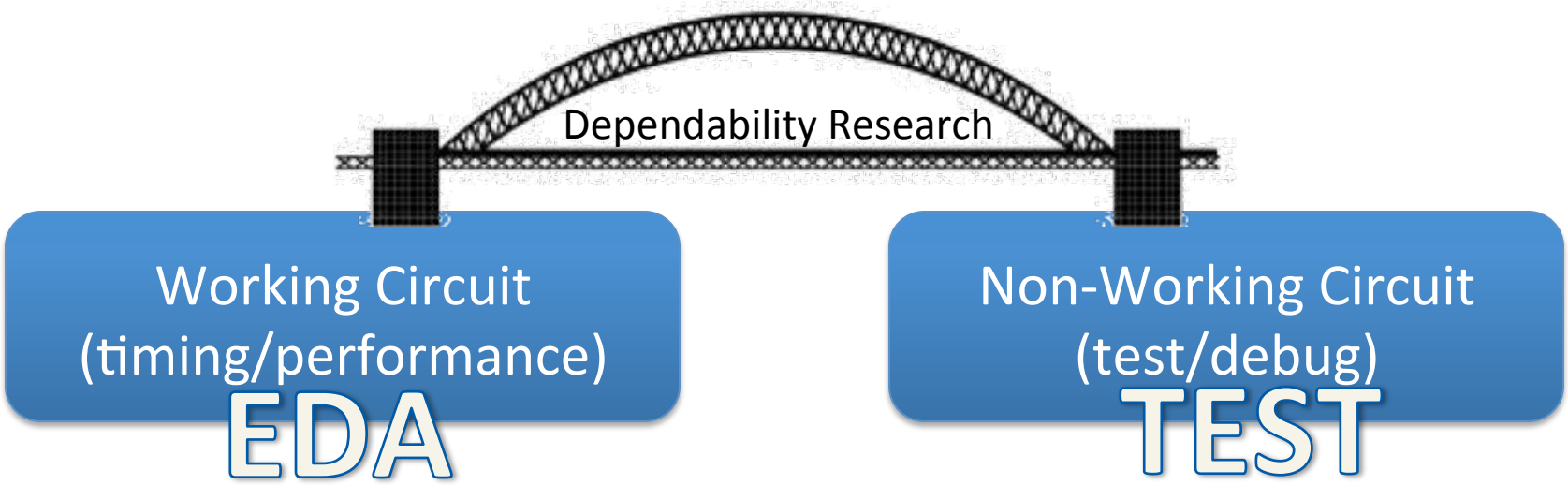
SRAM Yield vs. Corners

Dependability at the Circuit Level

- For challenging circuits like SRAM and Analog components, achieving dependability requires a robust design and analysis strategy (statistical capability +).
  - We can no longer afford to add margin unnecessarily.

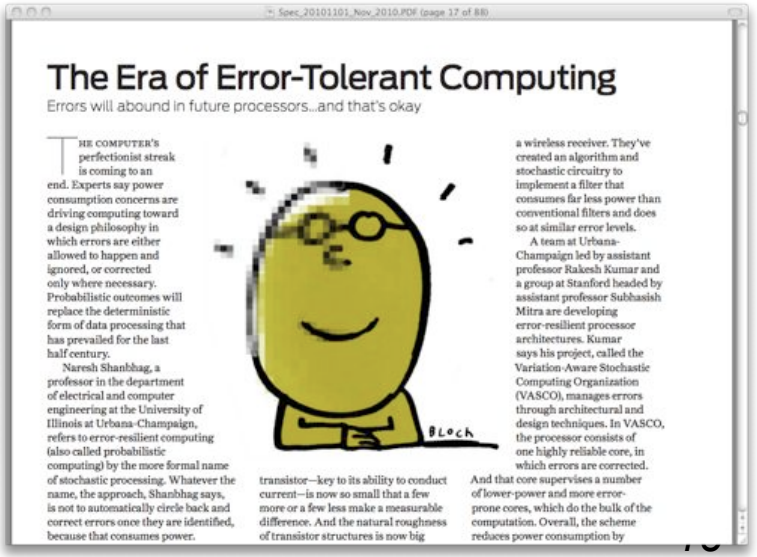
Observation:

- The bulk of the design tool chain assumes the circuit is "working" which is not appropriate for this type of research.



Conclusions

- We must keep Si/CMOS going for the foreseeable future.
  - This will require solving significant challenges in lithography, power and resilience... At the same time that R&D investment in this area is reducing because of consolidation.
- Circuits and devices appear to be increasingly failure prone and undependable with scaling.
  - For some components (e.g. SRAM) this has been recognized and substantially solved. For other circuits, work is needed.
- We must enable the standard EDA flow to naturally simulate & predict dependability.
  - Such enhancements will (finally) truly bridge the gap between design/test.
  - Target: allow future architects to accurately consider dependability as a 1st class requirement.



The Era of Error-Tolerant Computing

Errors will abound in future processors...and that's okay

THE COMPUTER'S perfectionist streak is coming to an end. Experts say power consumption concerns are driving computing toward a design philosophy in which errors are either allowed to happen and ignored, or corrected only where necessary. Probabilistic outcomes will replace the deterministic form of data processing that has prevailed for the last half century.

Naresh Shanbhag, a professor in the department of electrical and computer engineering at the University of Illinois at Urbana-Champaign, refers to error-resilient computing (also called probabilistic computing) by the more formal name of stochastic processing. Whatever the name, the approach, Shanbhag says, is not to automatically circle back and correct errors once they are identified, because that consumes power.

BLOCH

transistor—key to its ability to conduct current—is now so small that a few more or a few less make a measurable difference. And the natural roughness of transistor structures is now big

a wireless receiver. They've created an algorithm and stochastic circuitry to implement a filter that consumes far less power than conventional filters and does so at similar error levels.

A team at Urbana-Champaign led by assistant professor Rakesh Kumar and a group at Stanford headed by assistant professor Subhasish Mitra are developing error-resilient processor architectures. Kumar says his project, called the Variation-Aware Stochastic Computing Organization (VASCO), manages errors through architectural and design techniques. In VASCO, the processor consists of one highly reliable core, in which errors are corrected. And that core supervises a number of lower-power and more error-prone cores, which do the bulk of the computation. Overall, the scheme reduces power consumption by