Electrical & Computer ENGINEERING

ASPDAC 2007 Keynote

Next-Generation Design and EDA Challenges:

Small Physics, Big Systems, Tall Toolchains

Rob A. Rutenbar
Professor, Electrical & Computer Engineering
rutenbar@ece.cmu.edu

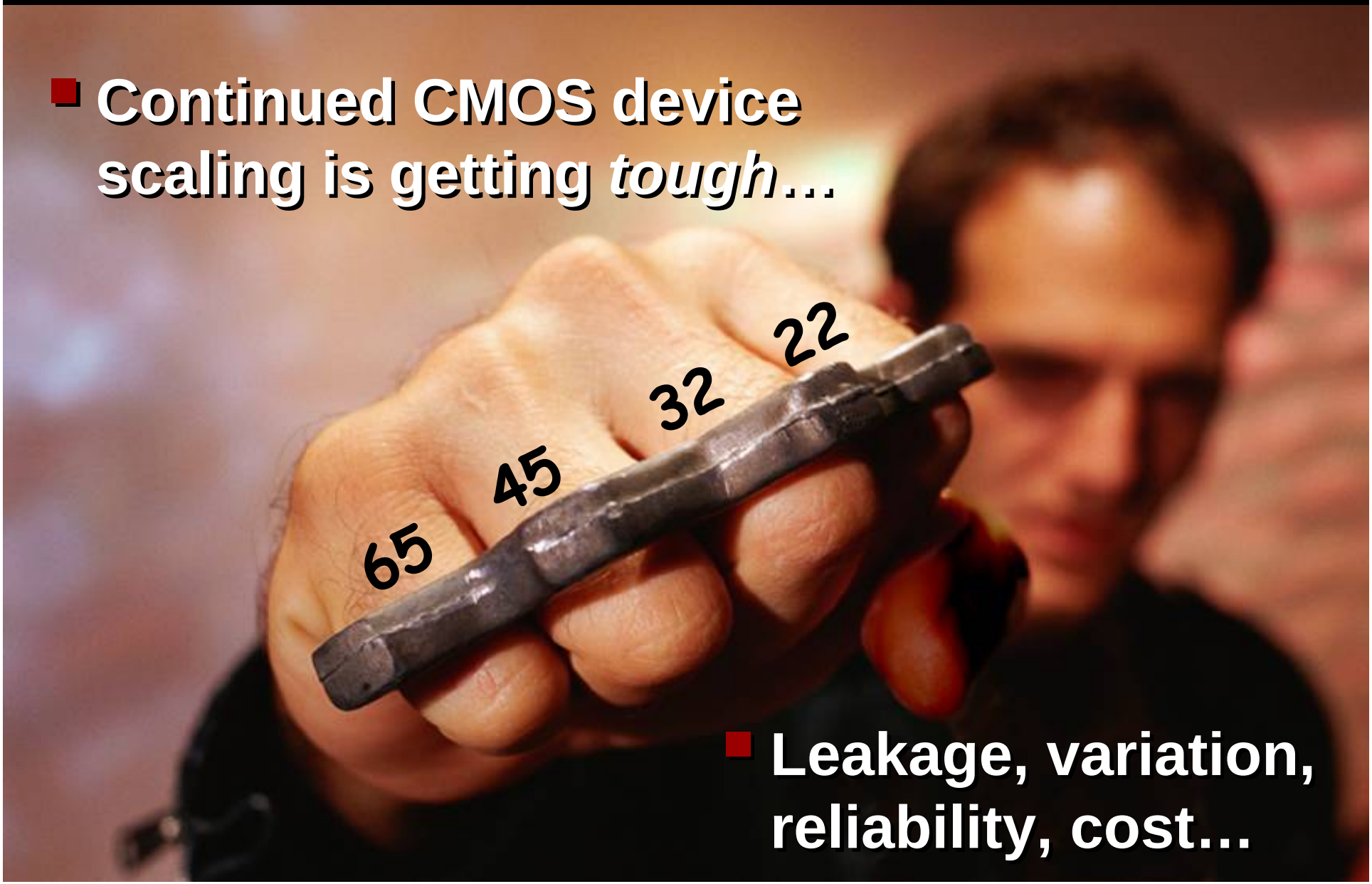
Carnegie Mellon

About This Talk...

- There are *three* different kinds of keynote talks...

#1: Explain A Big Problem

- **Continued CMOS device scaling is getting *tough*…**



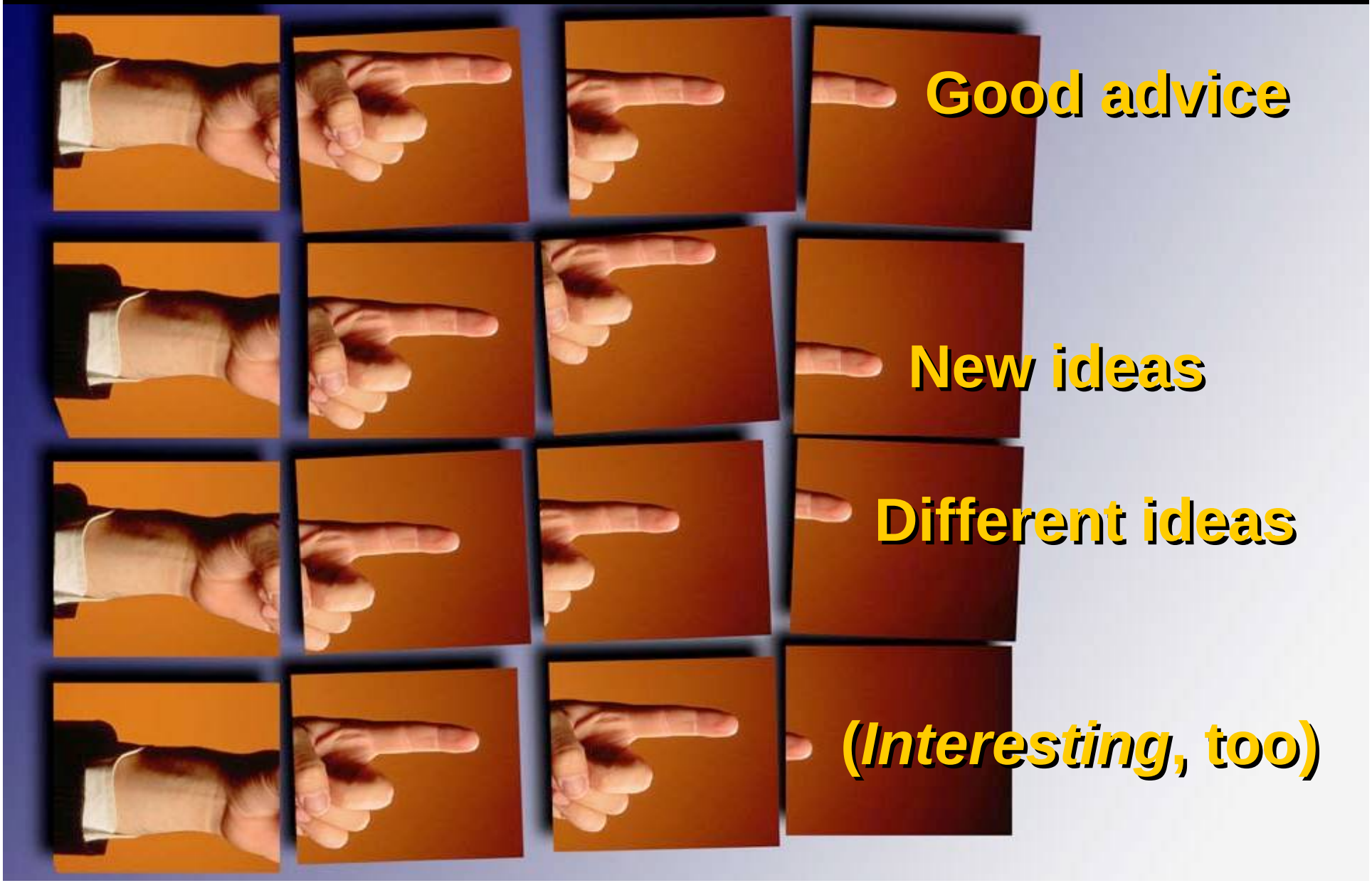
- **Leakage, variation, reliability, cost…**

#2: Predict the Future

- Fun...

...but *difficult*

#3: Offer Some Advice

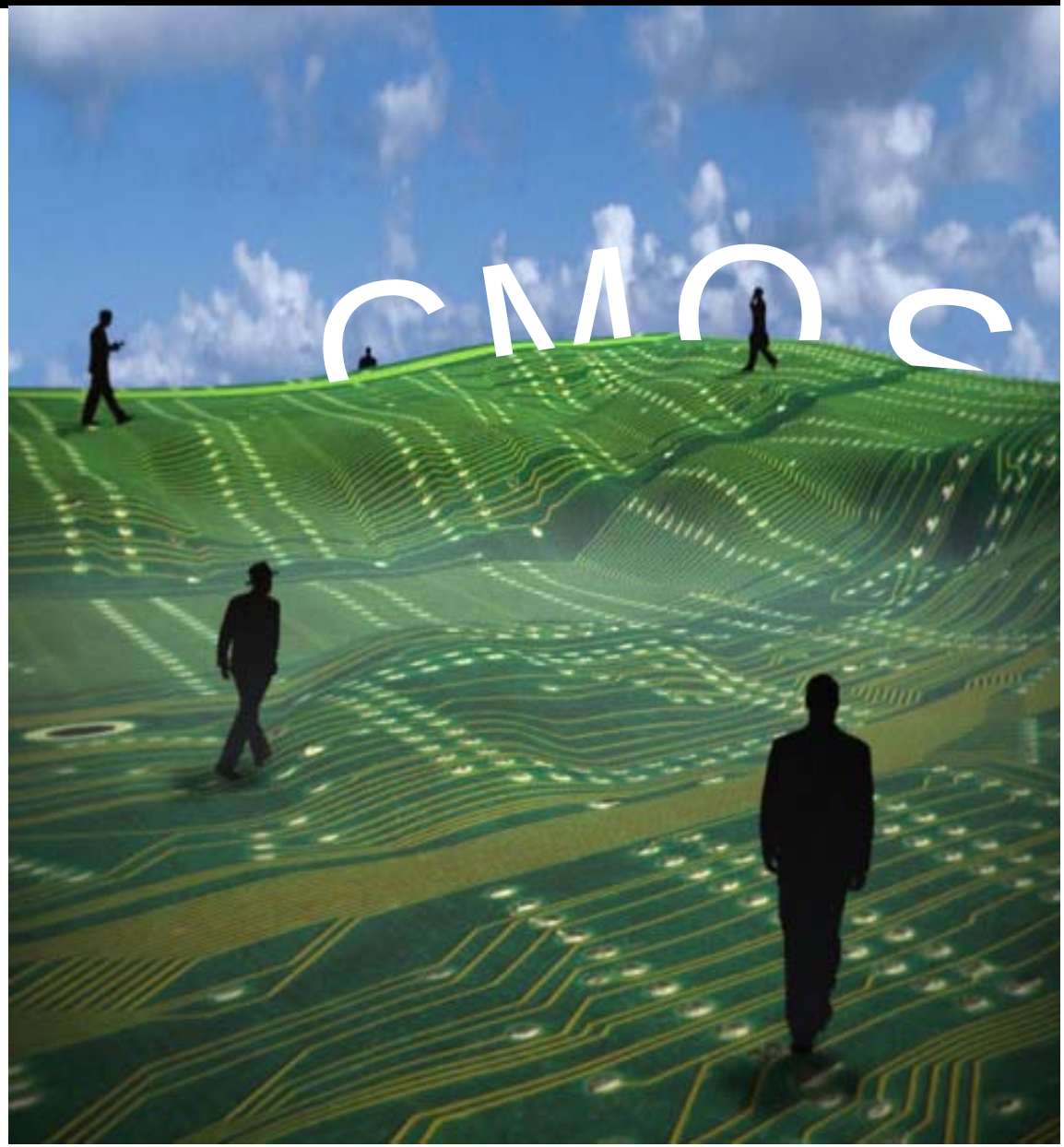


This Is An *Advice* Talk

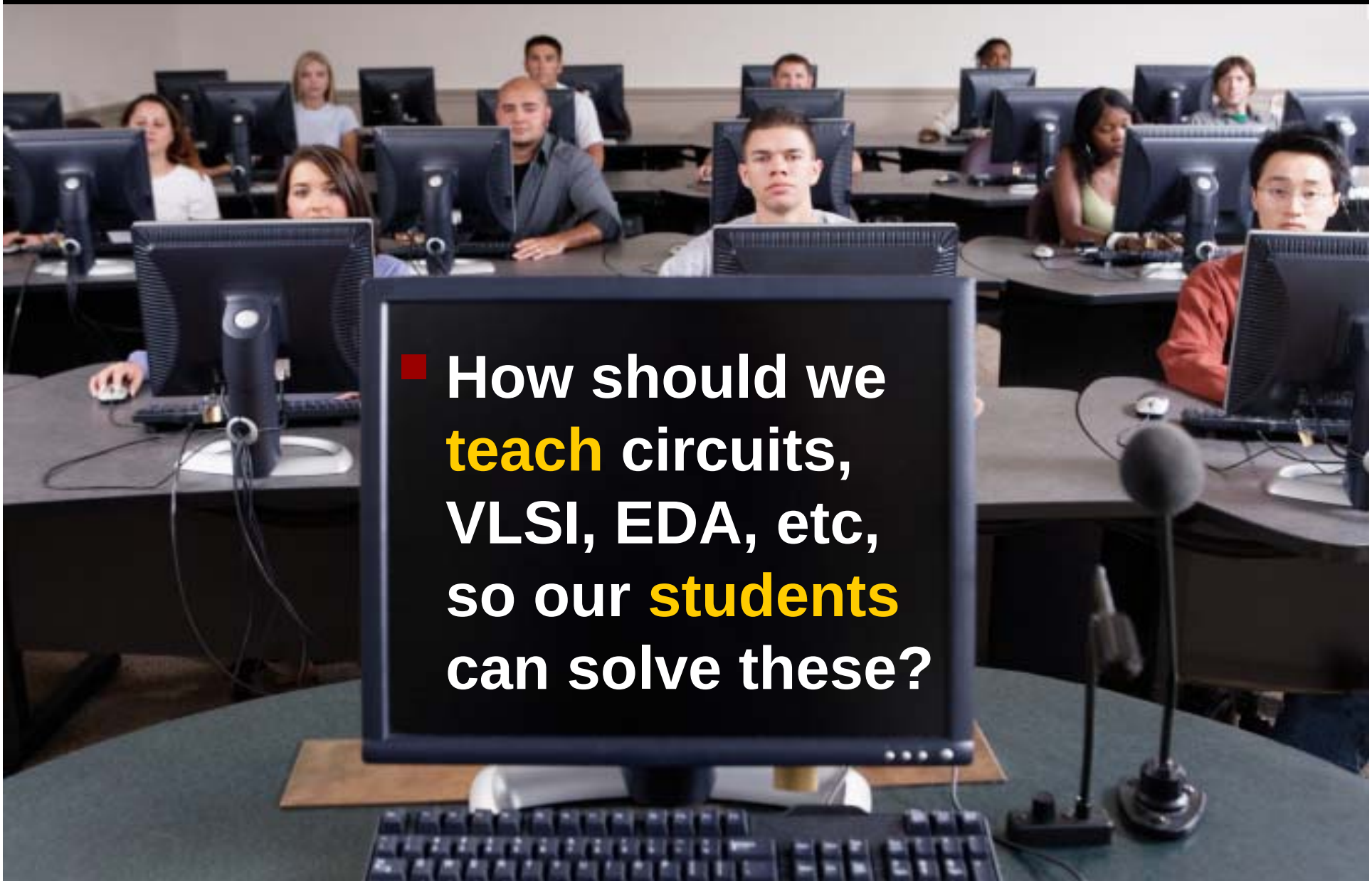
As we near the end of the silicon roadmap...

What *kinds* of problems are the big challenges?

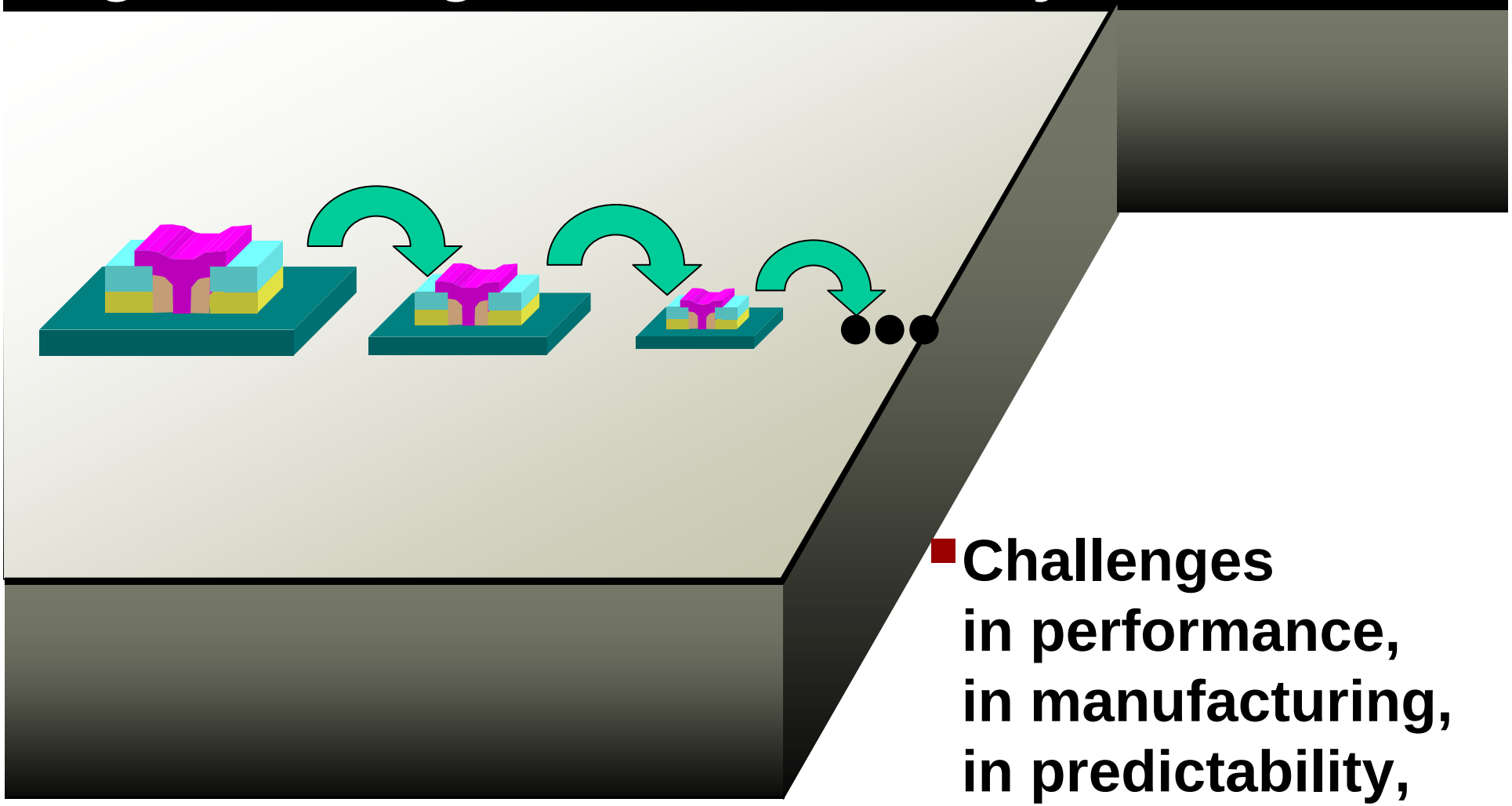
What *styles* of problem-solving might work best?



Secondary, Personal Motivation

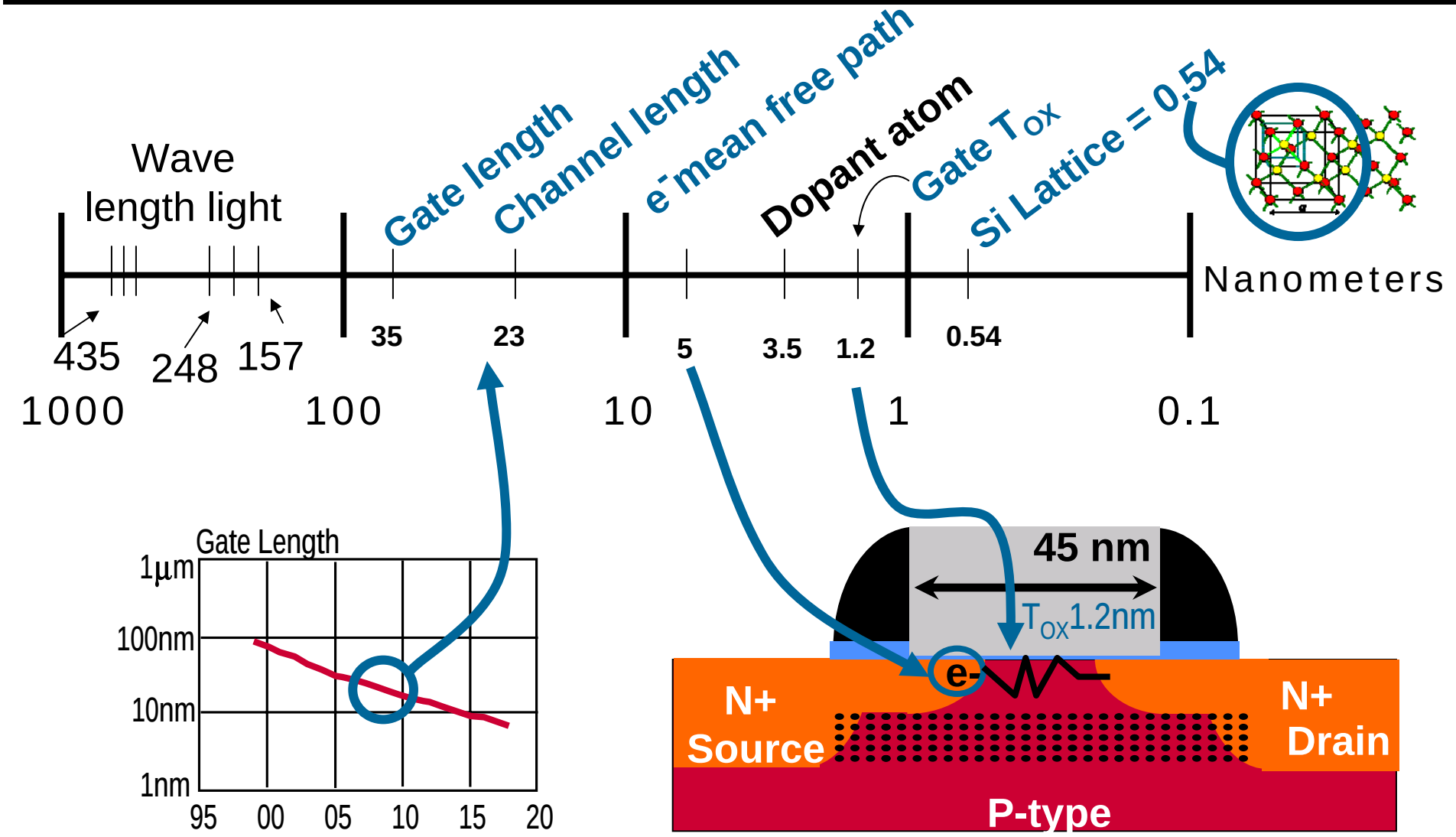


Big Challenge #1: Small Physics



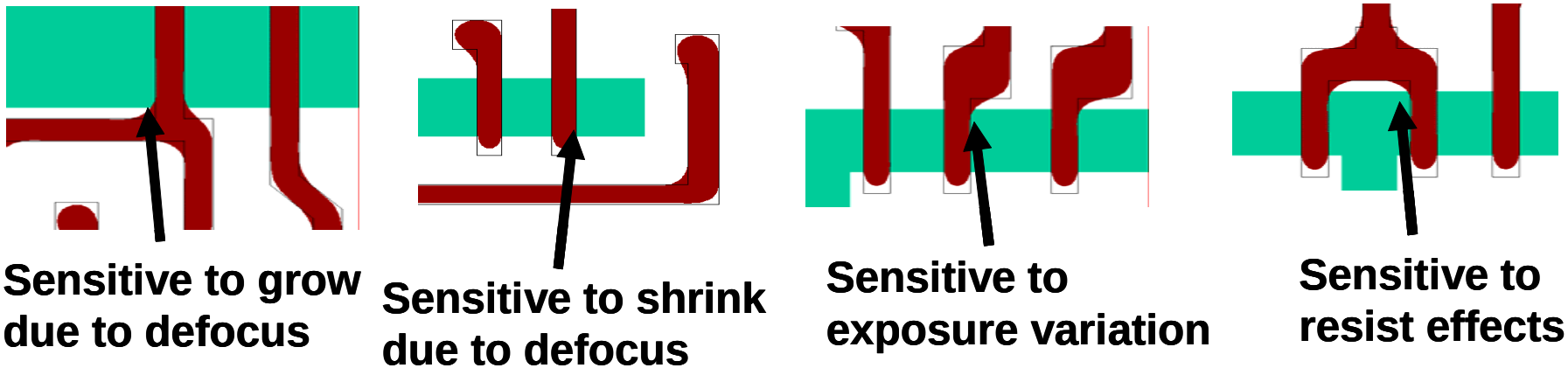
- **Challenges in performance, in manufacturing, in predictability, in cost...**

Approach *Atomic* Scale → *Challenges*



[Source: Scott Thompson, U Florida]

One Challenge: Mask Variability & Cost

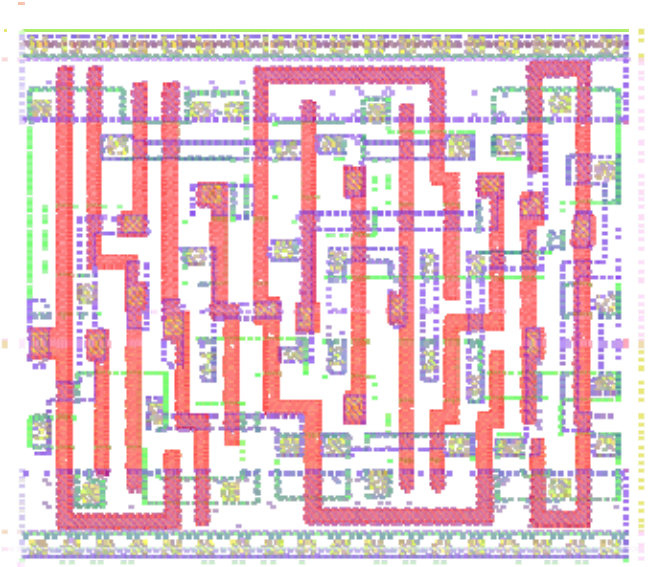


- **Standard geometry rules may *not* be enough**
  - **DRC rules *very* complex; get *worse* with scaling**
  - **Sub-wavelength lithography; neighbor interactions**

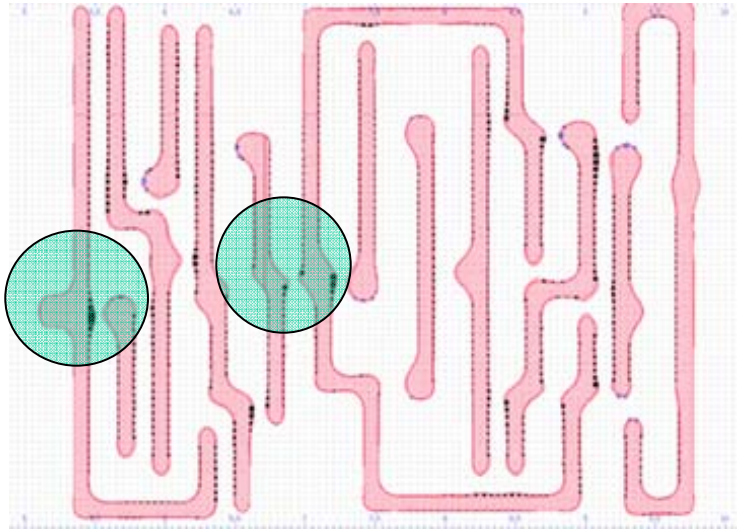
[Source: Larry Pileggi, Andrzej Strojwas, CMU]

The Mask Variability/Cost Challenge

- **Big problems on even simple circuits**
  - **Performance (e.g., leakage), yield compromised**



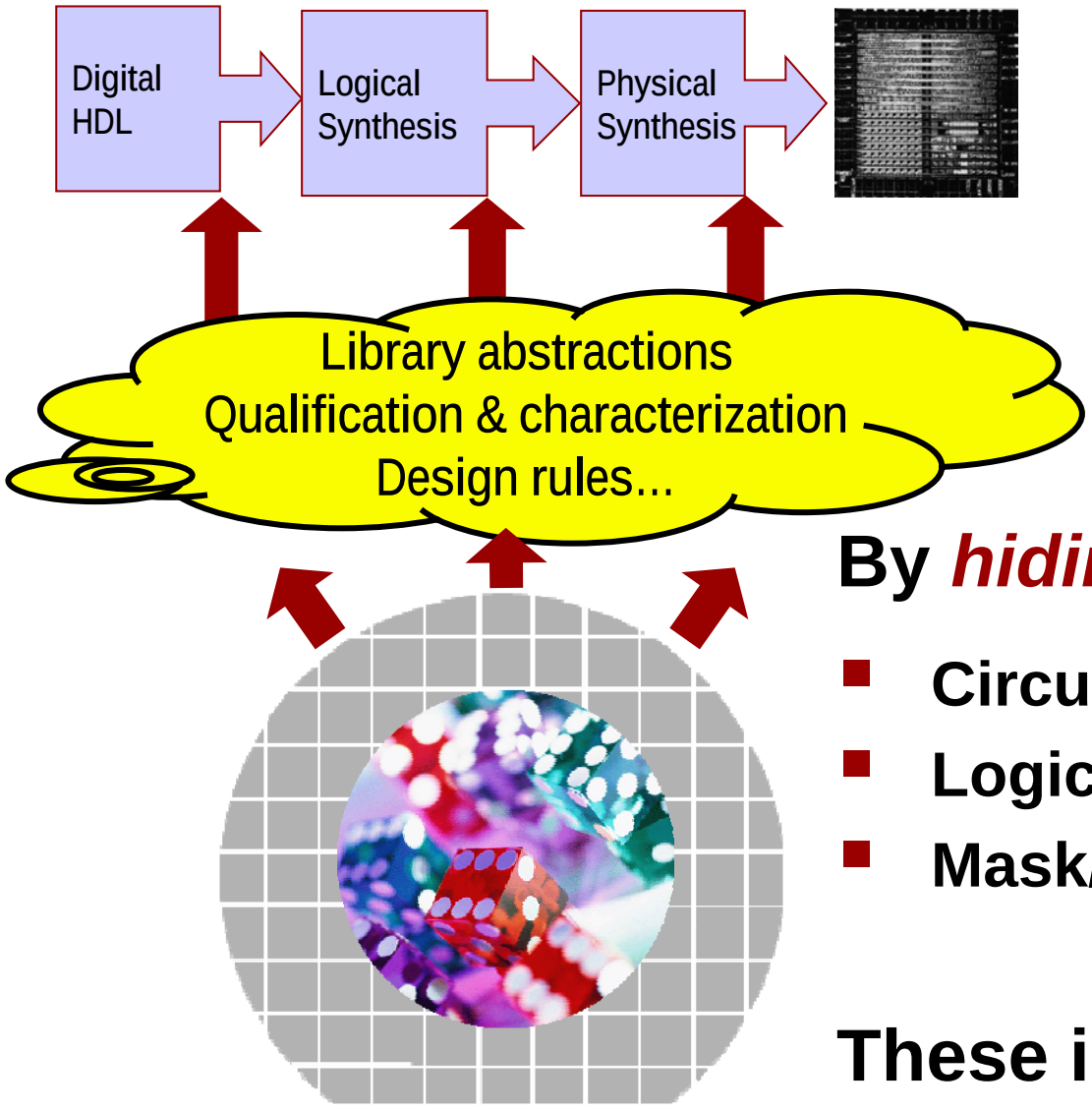
Typical standard cell



Printability problems from lithography simulation

[Source: Larry Pileggi, Andrzej Strojwas, CMU]

How Did We Handle This Until Now…?



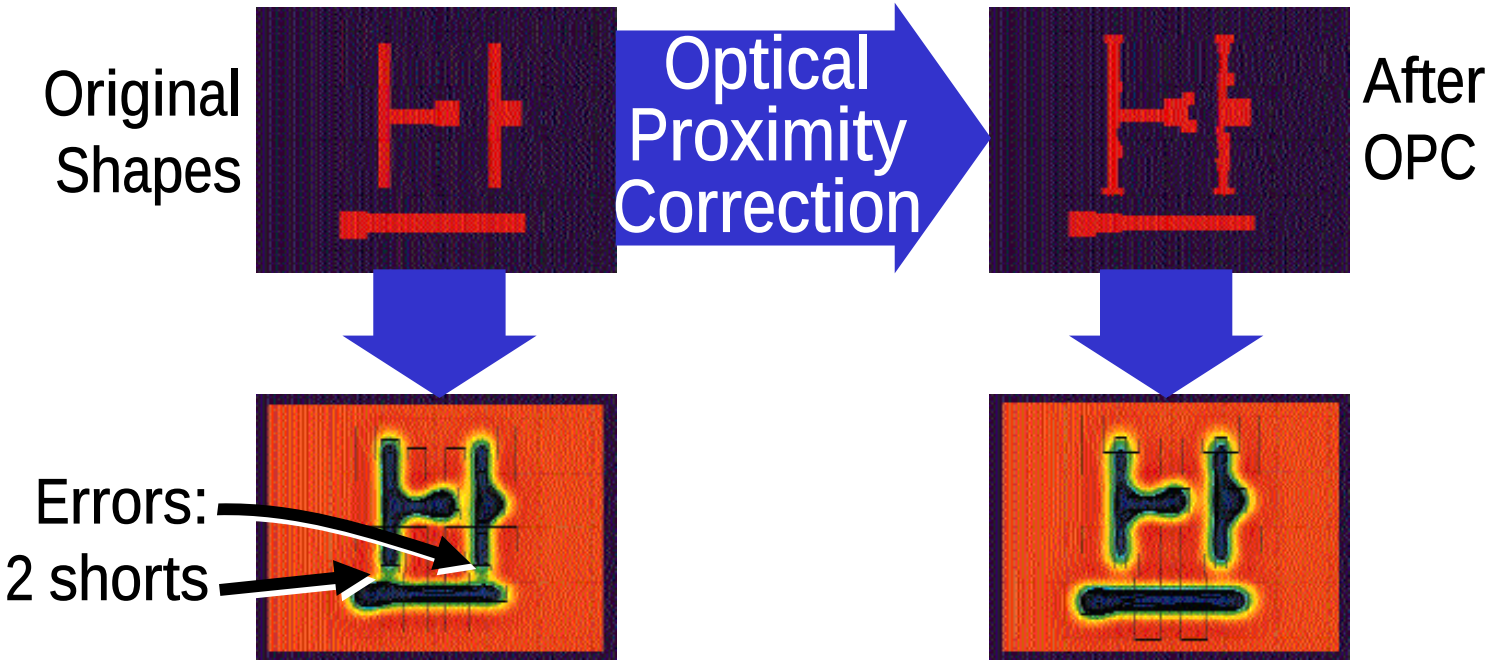
By *hiding & abstracting*

- **Circuit & shapes rules**
- **Logic/memory libraries**
- **Mask/circuit/logic isolated**

These ideas starting to *fail*

How to Explain (*Teach*) This...?

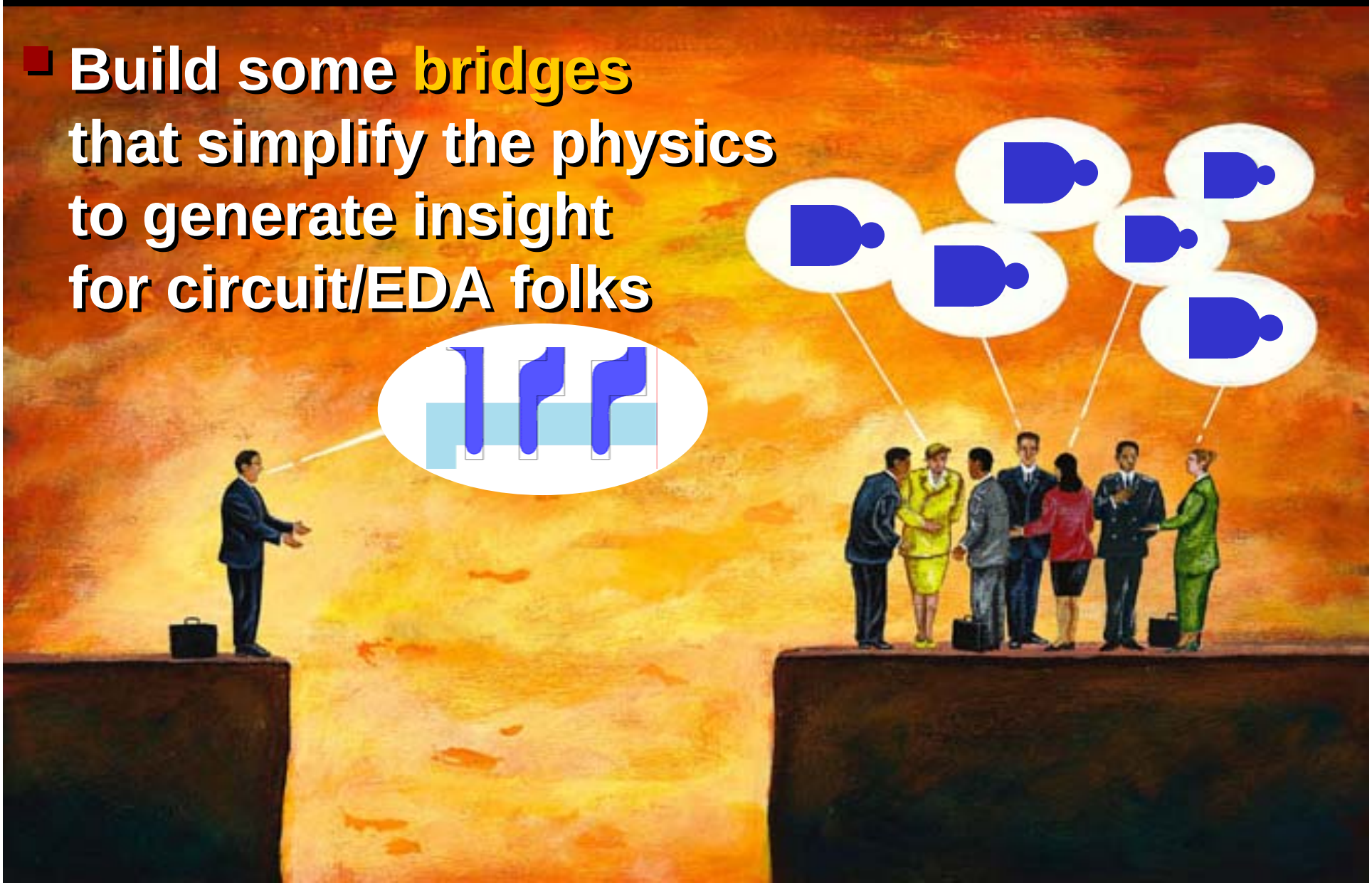
- **Before nanoscale CMOS, masks were simple**
  - **Today, cannot make a mask without complex *Resolution Enhancement Technology* (RETs)**



[www.isdmag.com/Editorial/1999/coverstory9905.html]

“Small Physics” Advice

- **Build some bridges that simplify the physics to generate insight for circuit/EDA folks**



Some Useful Insight...

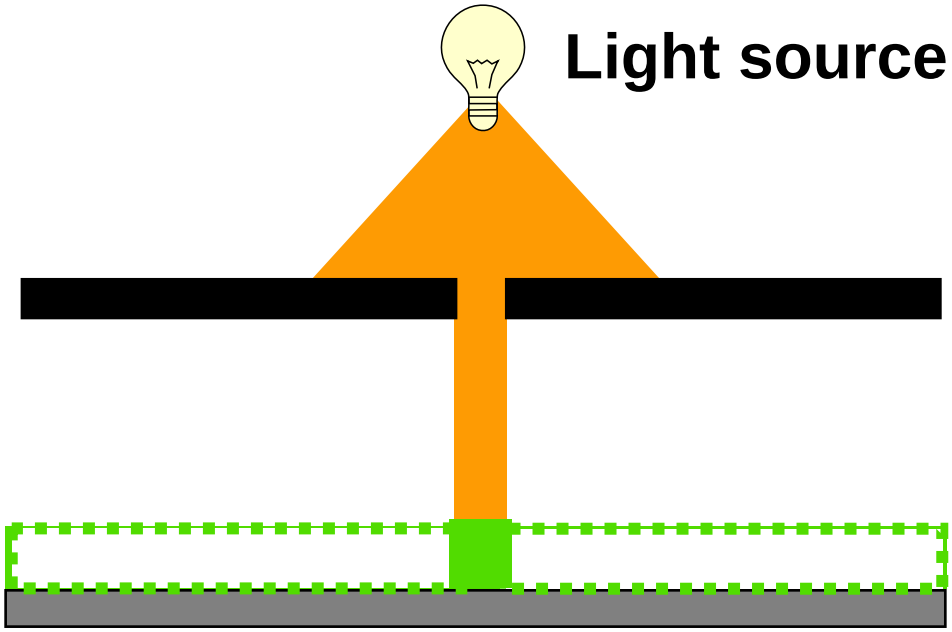
"One thing we know about creativity is that it typically occurs when people who have mastered two or more quite different fields use the framework in one to think afresh about the other..."

Marc Tucker, *Tough Choices or Tough Times*

- So: Lithography + circuits/CAD/optimization?

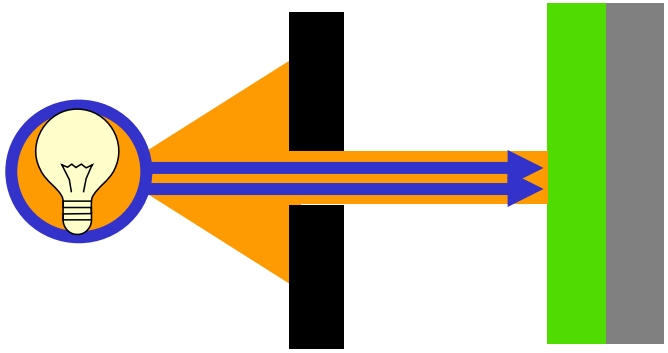
Example: Vastly Simplified Litho Model

- **Light source + 1D mask + photoresist**

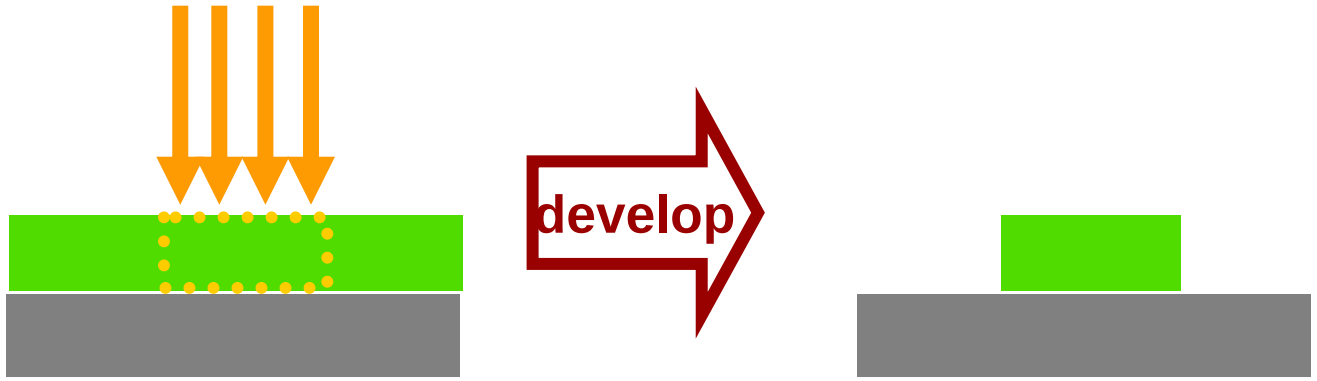


Unpleasant Physical Realities

- **Light does *not* travel in nice straight lines**

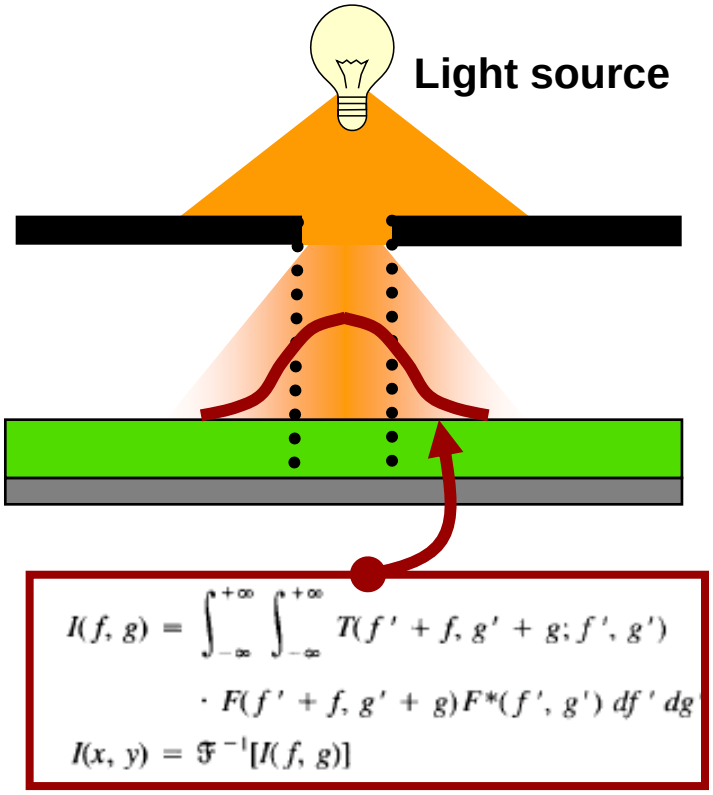


- **Resist *not* infinitely *sensitive* to light**

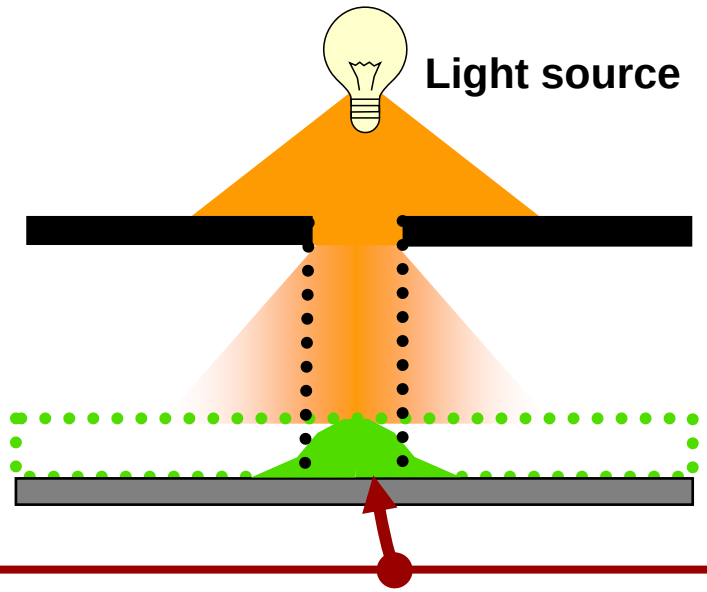


Light *Bends* and Resist is *Nonlinear*

Light intensity delivered is complex function of litho optics

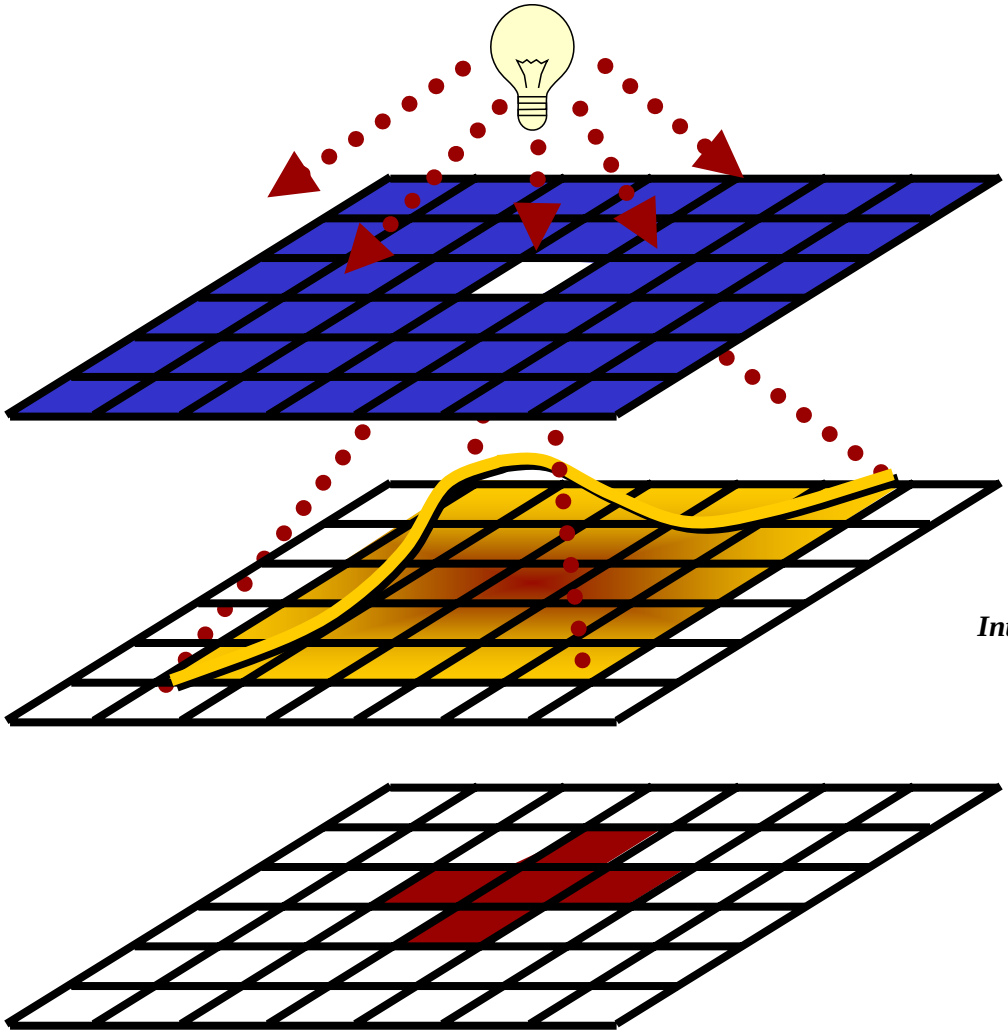


Resist response is complex function of light delivered



Complex optics + chemistry = Complex mask shape

To Get Some Insight: *Simplify*

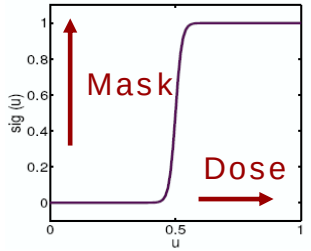


Pixeled binary mask

Additive Gaussian blur optics

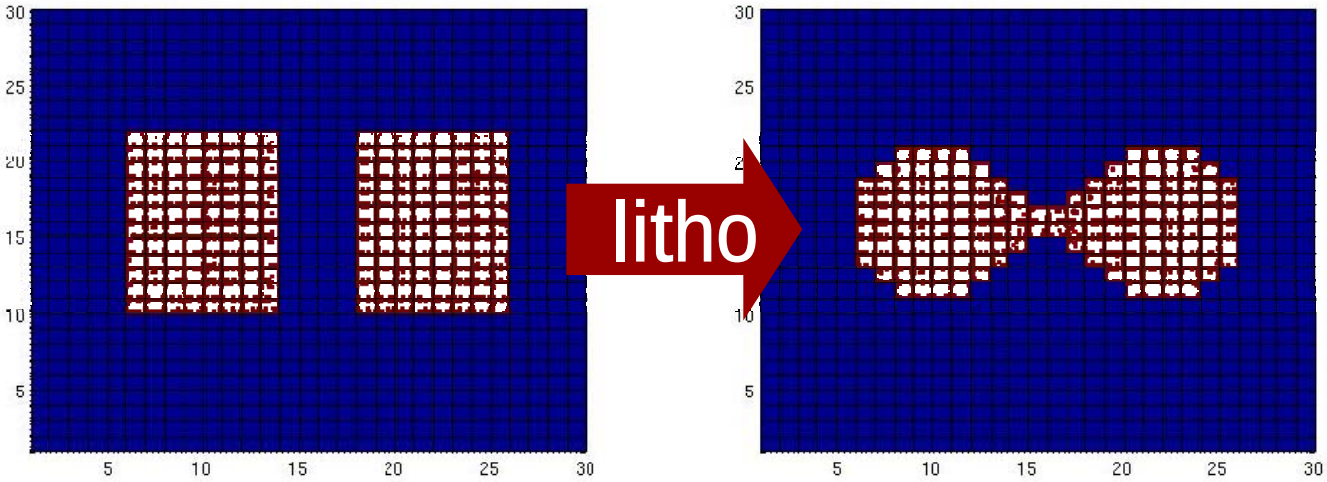
$$Intensity(\Delta x, \Delta y) = N \bullet \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Delta x)^2}{2\sigma^2}} \bullet \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Delta x)^2}{2\sigma^2}} \right]$$

Simple resist threshold



A Nice CAD Class Project: OPC

- **Optical Proximity Correction**
  - ***Inverse mask* problem: given shapes, *derive* mask**
  - **Turn pixels on/off with simple optimizer (inspired by early OPC+annealing ideas of Zakhor et al, Berkeley)**

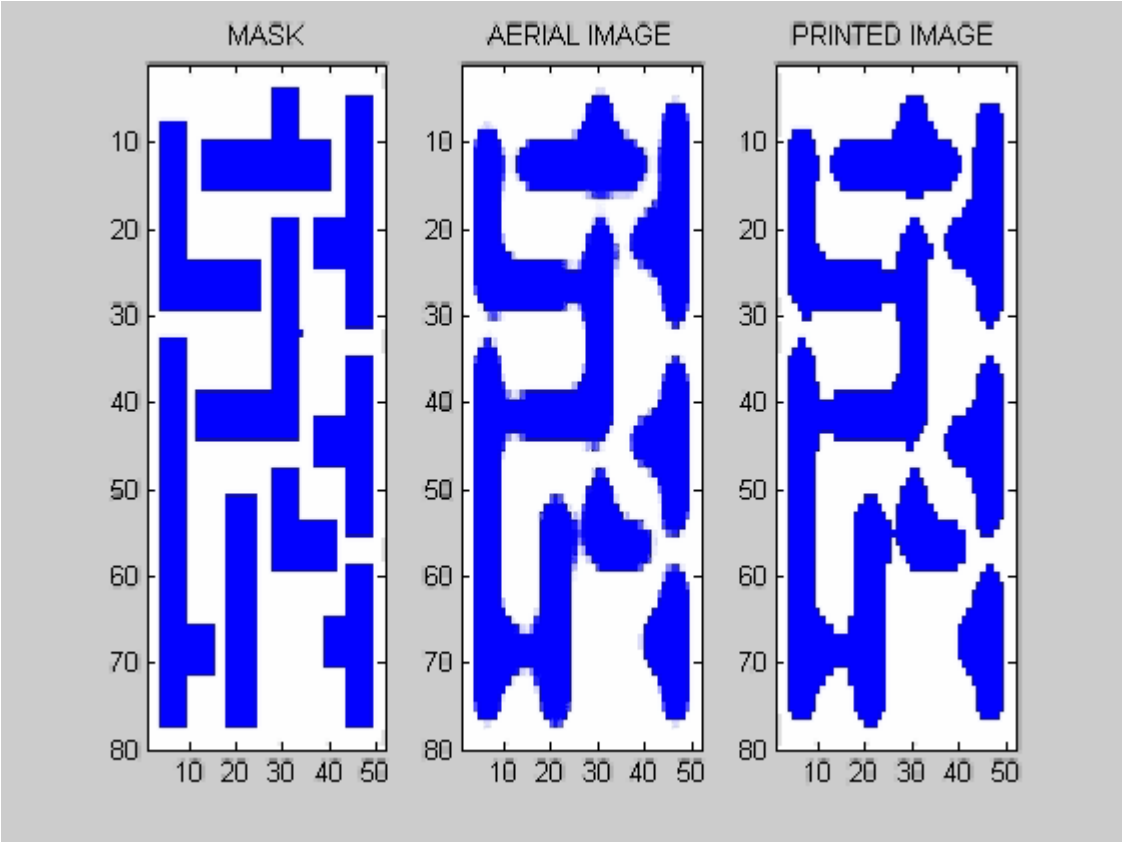


Cannot just use desired shapes

"Fake" physics model gives realistic aerial image

Even with “Fake” Physics: Useful

- **Excellent litho insights for circuit/EDA people**



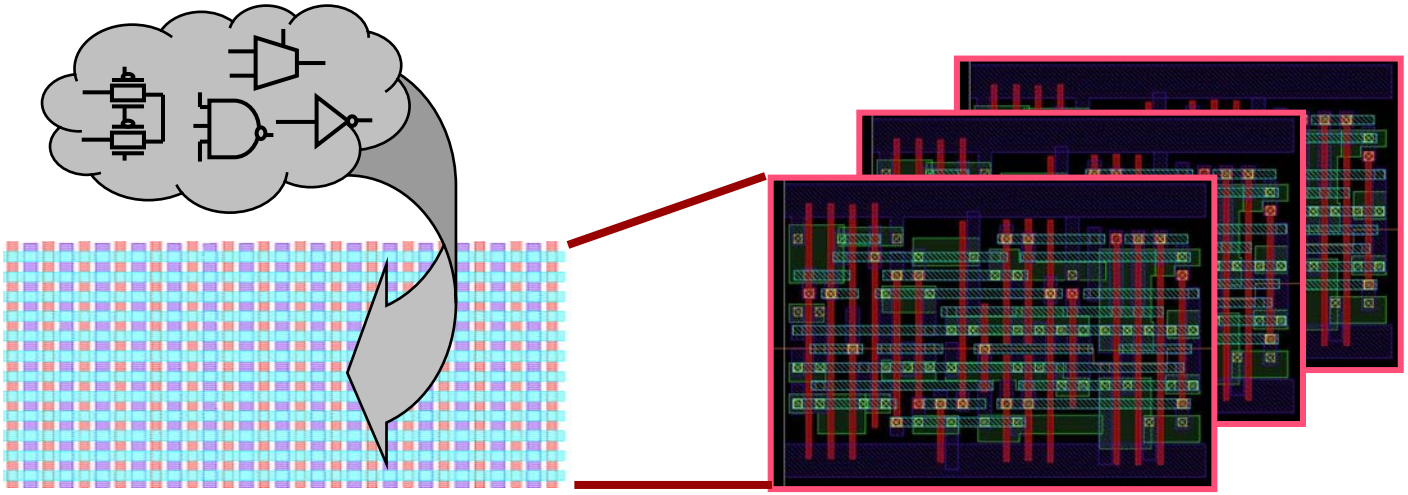
[Source: Sonia Singhal, CMU]

Why This Matters: Regular Circuits

- Next generation of designs may rely on a totally new idea: regular circuit fabrics, where all transistors ~identical

Example: CMU Regular Logic "Bricks"

- **Small, dynamically compiled, *litho-regular*, configurable logic lib, few geometry patterns**
- **Design flow requires novel tools, algorithms**

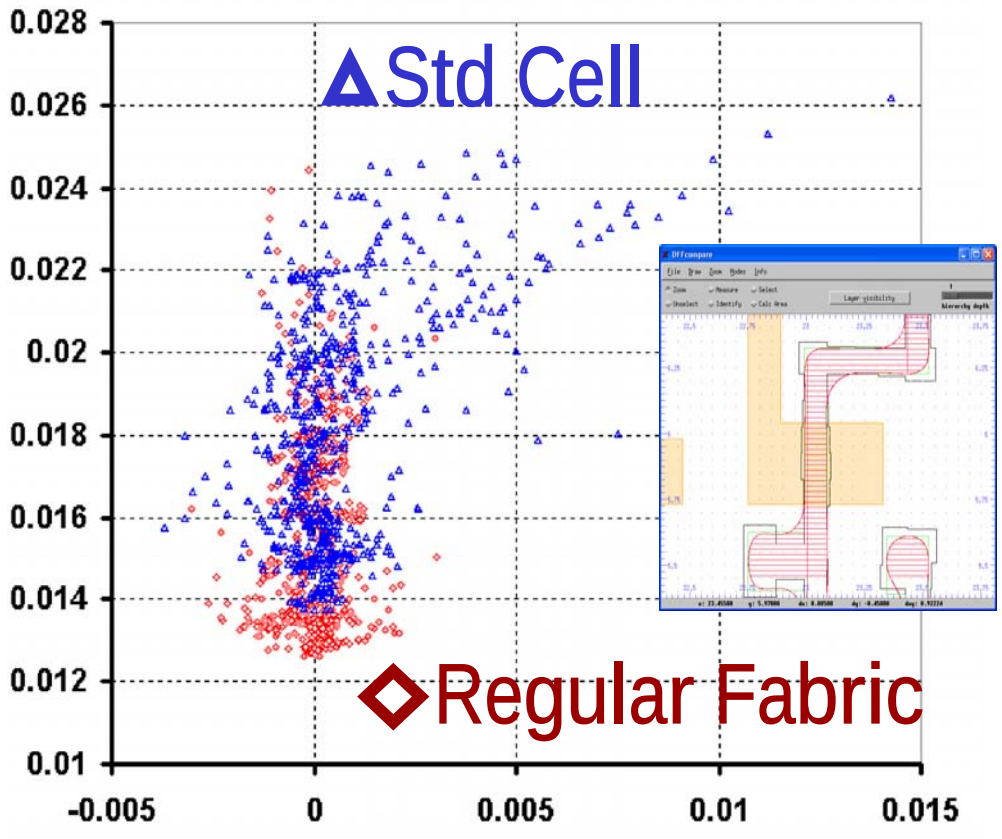


[Source: Larry Pileggi, Andrzej Strojwas CMU]

Early Results from Regular Bricks

- **Dramatically better L_{eff} control → Lower leakage**

ΔL_{poly} [µm]
printed (nom) -
printed (w/ defocus)

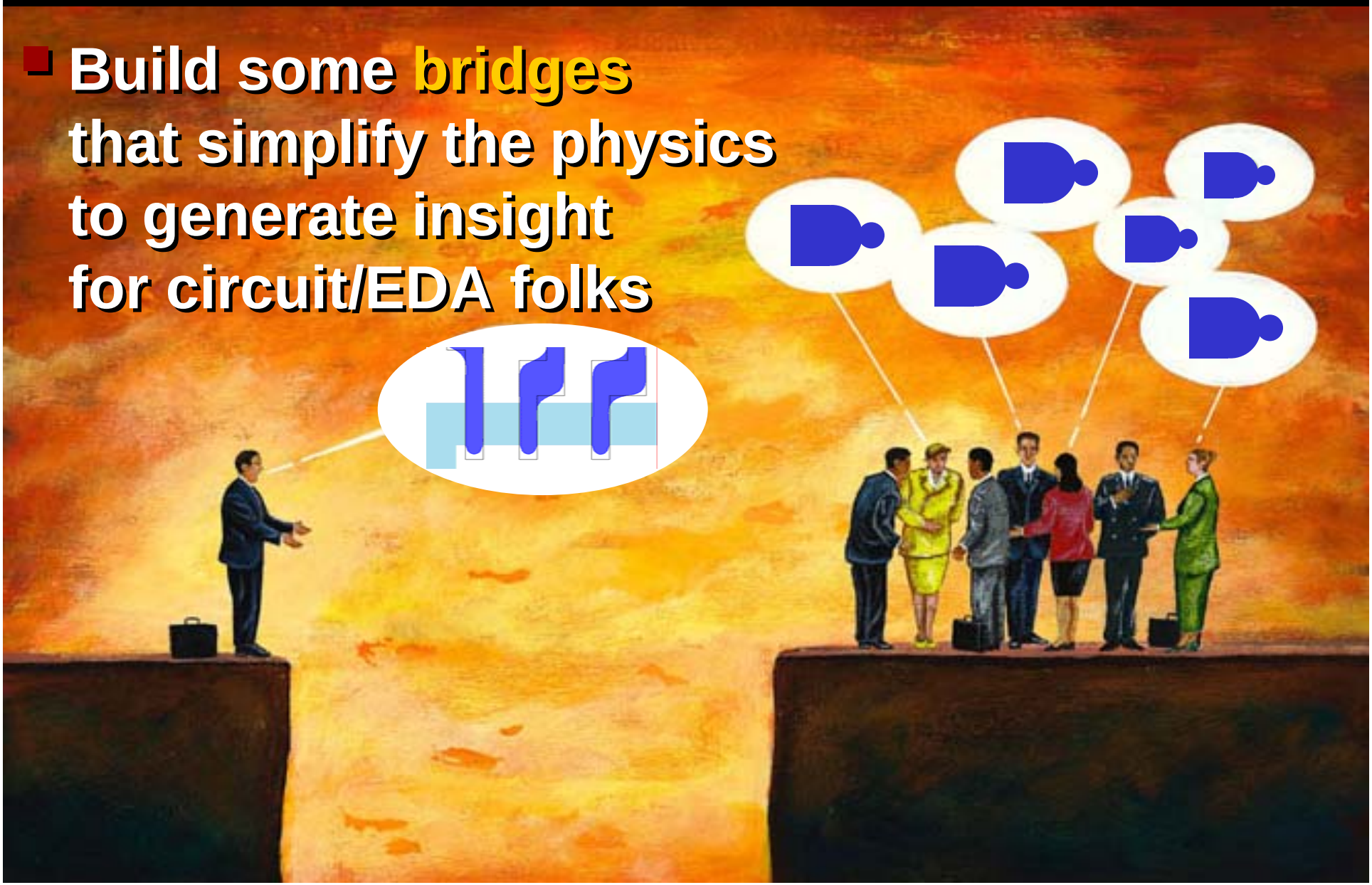


ΔL_{poly} [µm]
drawn (target) - printed (nom)

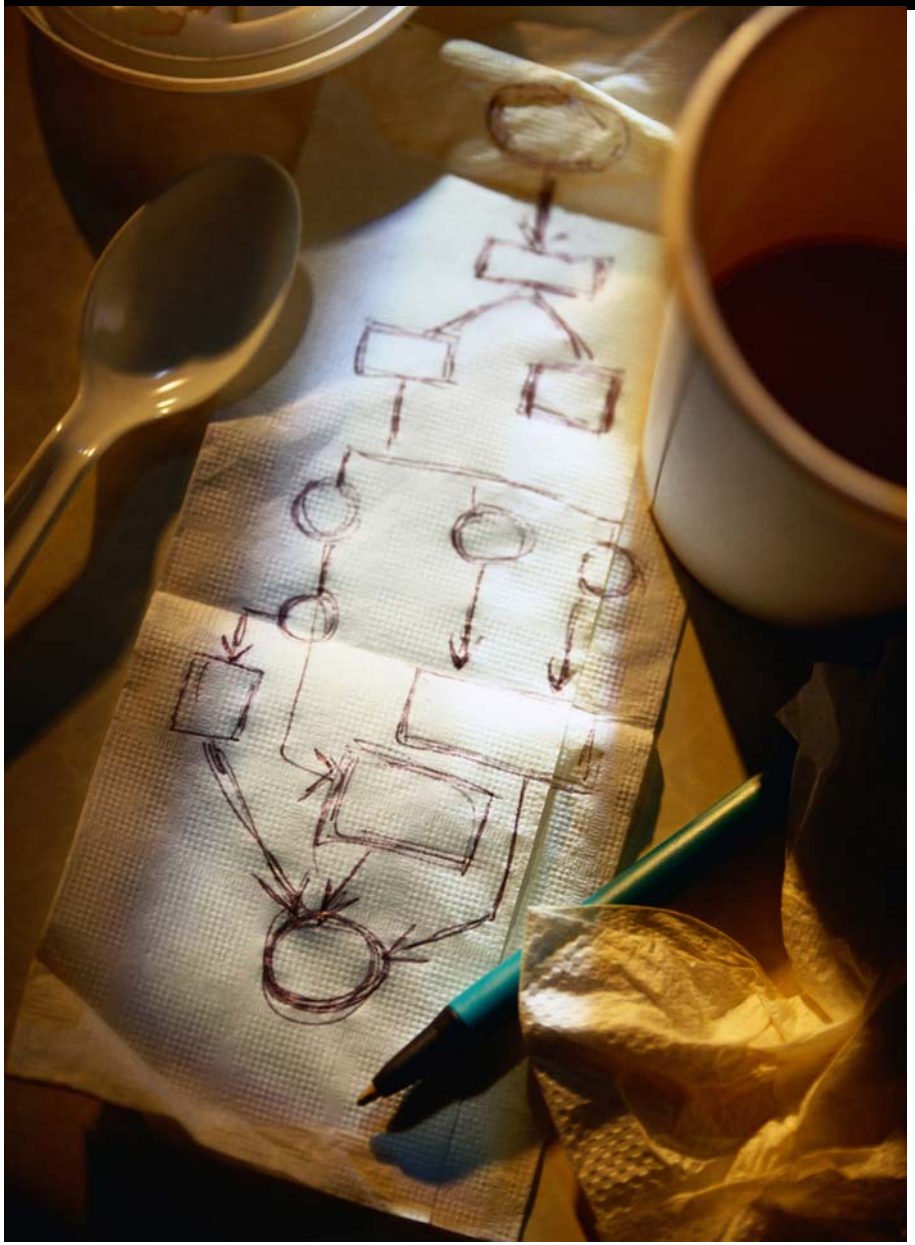
[Source: Pileggi/Strojwas, CMU]

“Small Physics” Advice

- **Build some bridges that simplify the physics to generate insight for circuit/EDA folks**



Big Challenge #2: Tall Tool-Chains

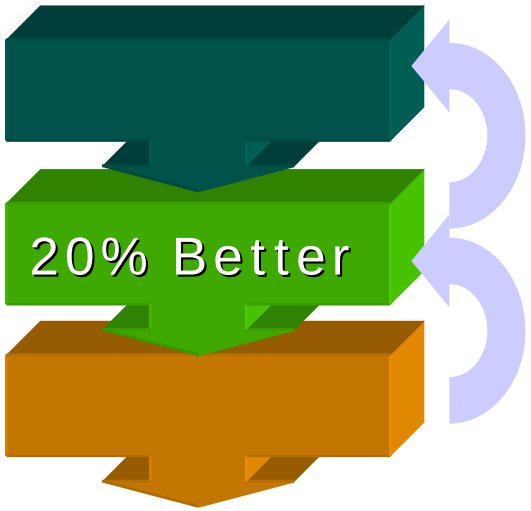


The sets of CAD tools we use to do big designs are *not simple* anymore

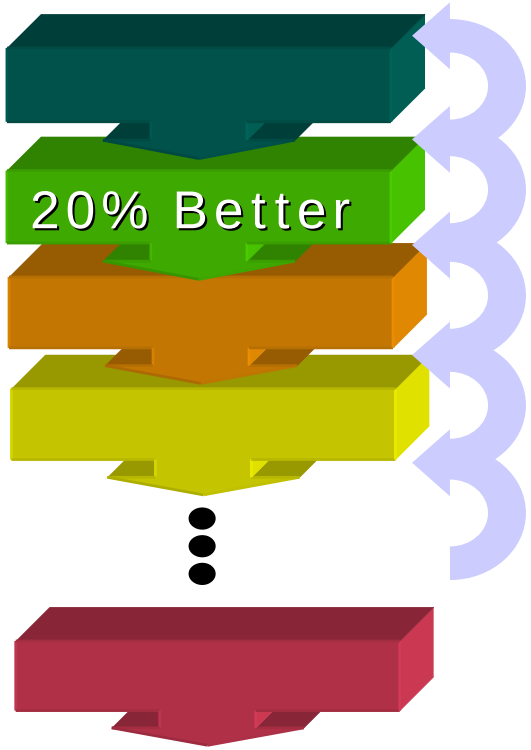
This complexity poses new problems for *tool innovation*

Rutenbar's *Rule of Attenuation*

- **The taller the tool-chain, the more difficult for innovation in 1 tool to "survive the flow"**

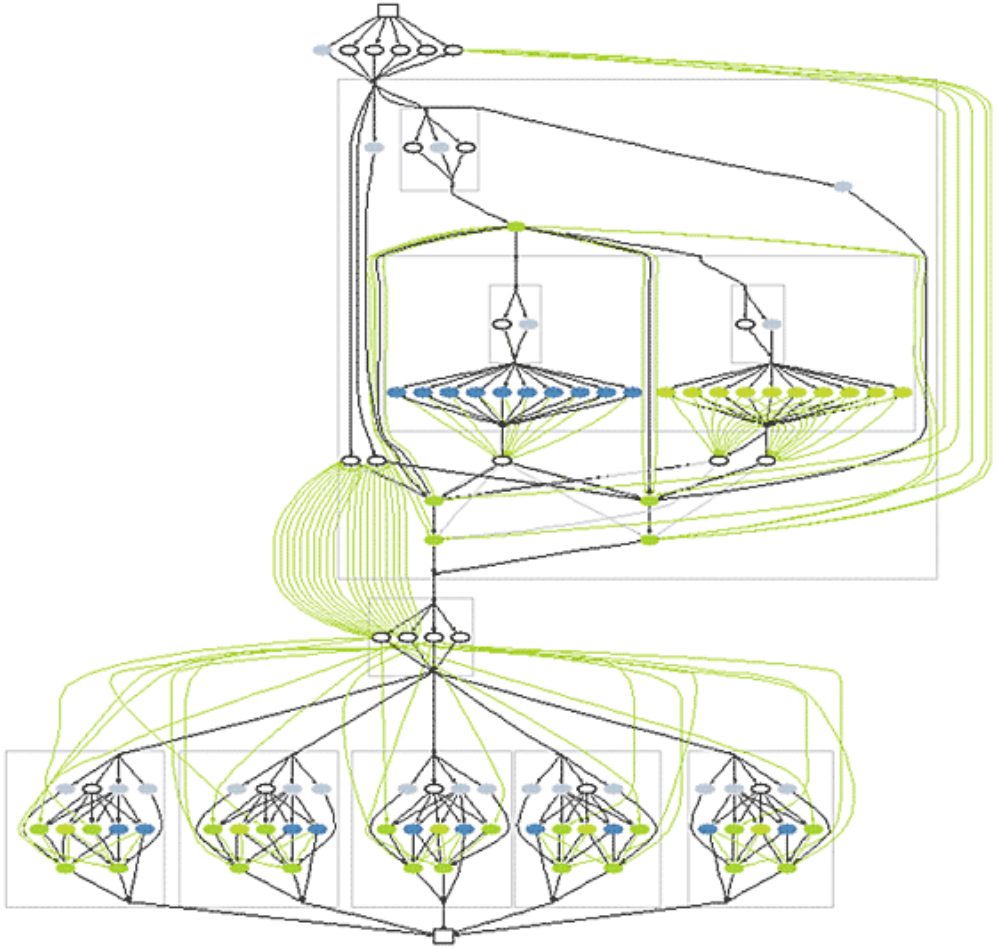


~10% Better



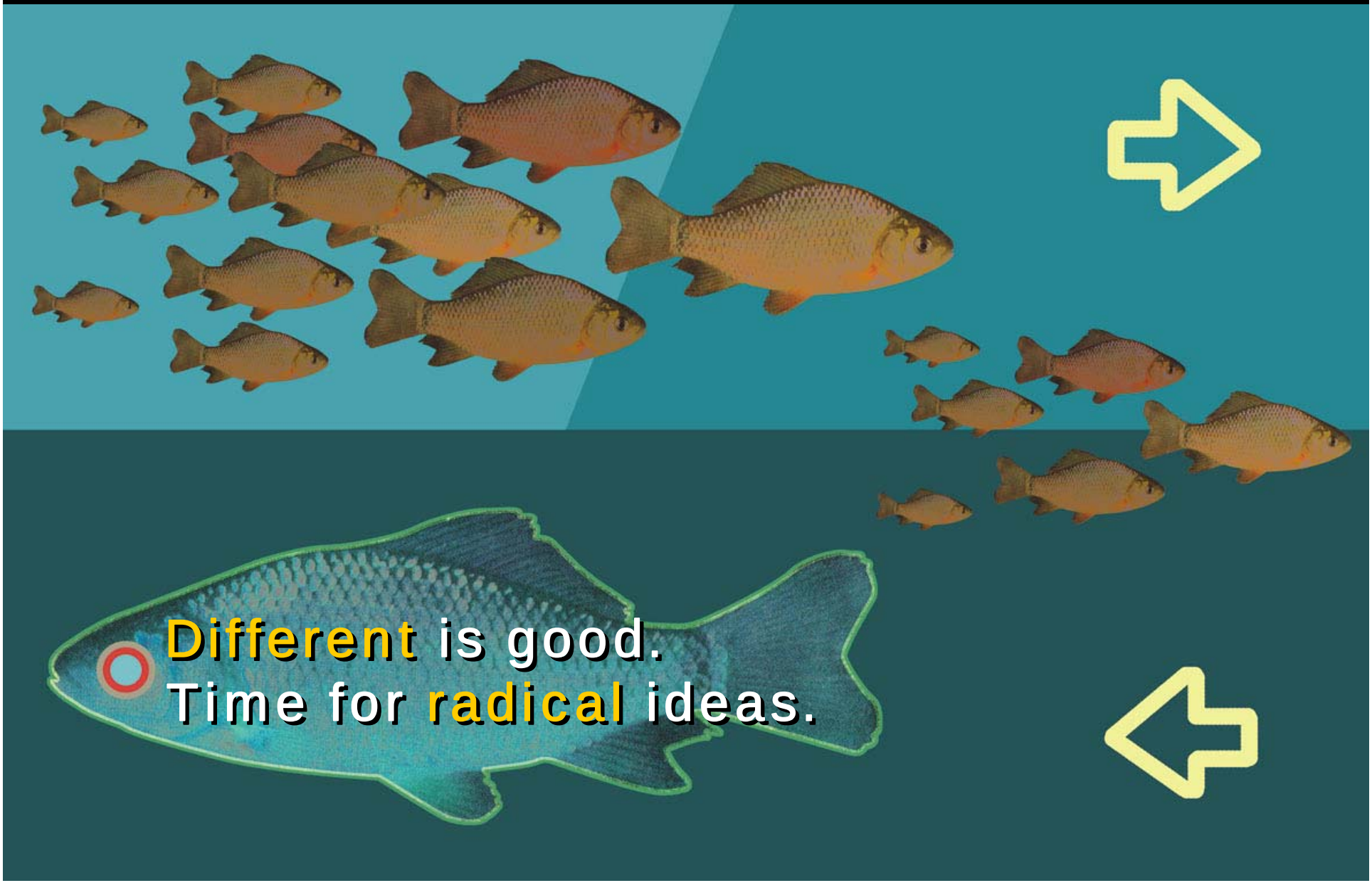
< 1% Better?

How “Tall” is Tall?



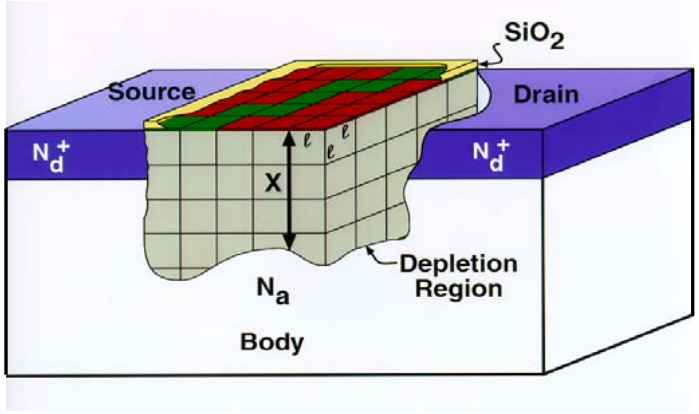
- ***Very* tall**
- ***Very* deep sets of connected tools, scripts, files, databases, sign-offs, etc**
- **Ex: 4000-step commercial design-flow**

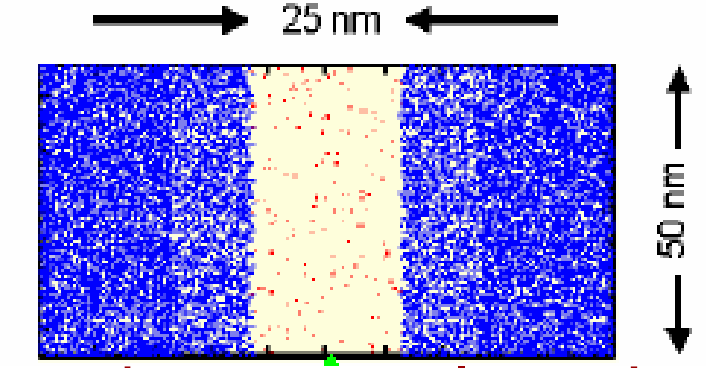
“Tall Tool-Chains” Advice



Back to a "Small Physics" Problem

- **At nanoscale, *random disturbances* have a significant impact on devices**

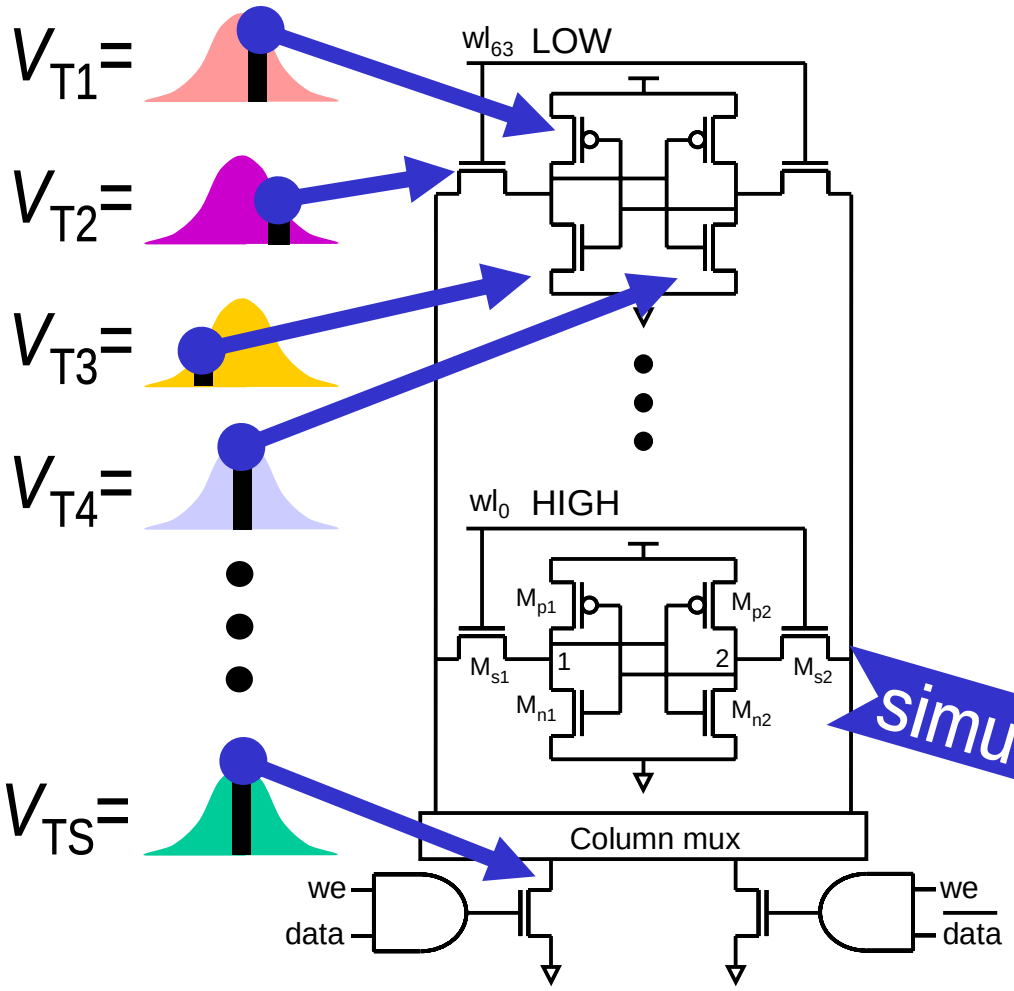




Random Dopant Fluctuation

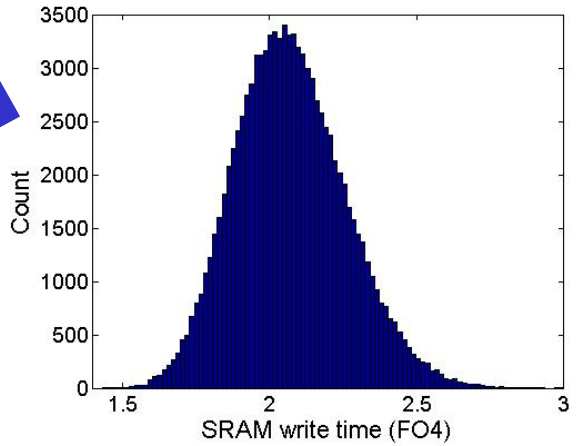
- **How significant?**
  - **25nm channels, $\sigma[V_T]$ uncertainty → 10-20%**
  - **From [Wong, 1999 VLSI Tech. Symposium]**

To Evaluate Circuit Impact: Monte Carlo



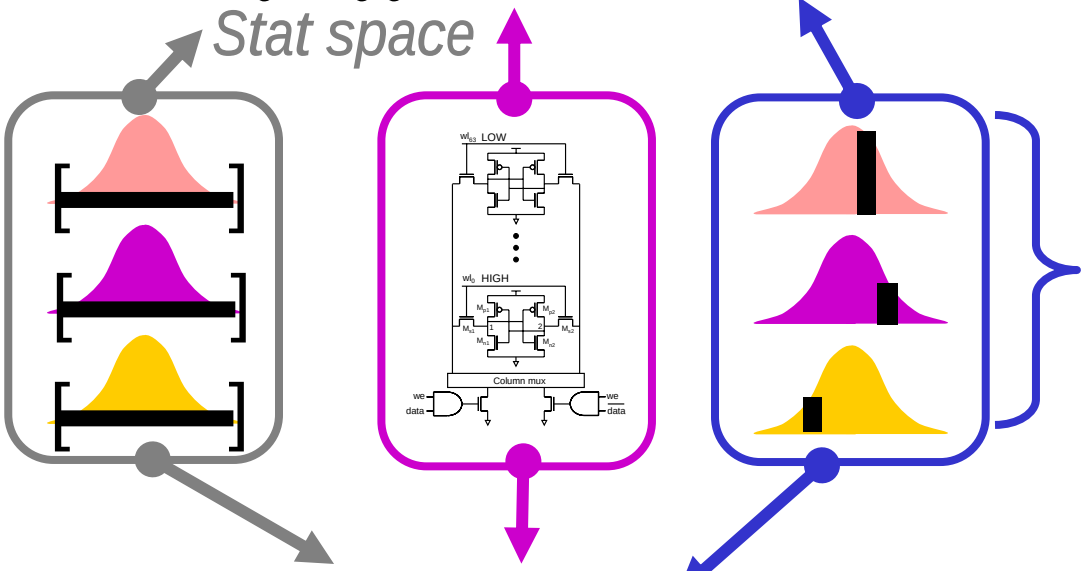
simulate

- **Sample each statistical variable**
- **Parameterize one circuit, simulate it**
- **Repeat--n samples**



Monte Carlo Math: Just A Big Integral

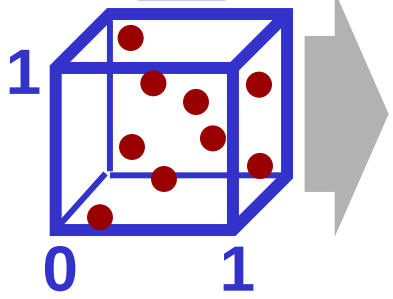
$$\text{Est} = \int \cdots \iint f_{ckt}(x)\, p(x)\, dx$$



s statistical vars

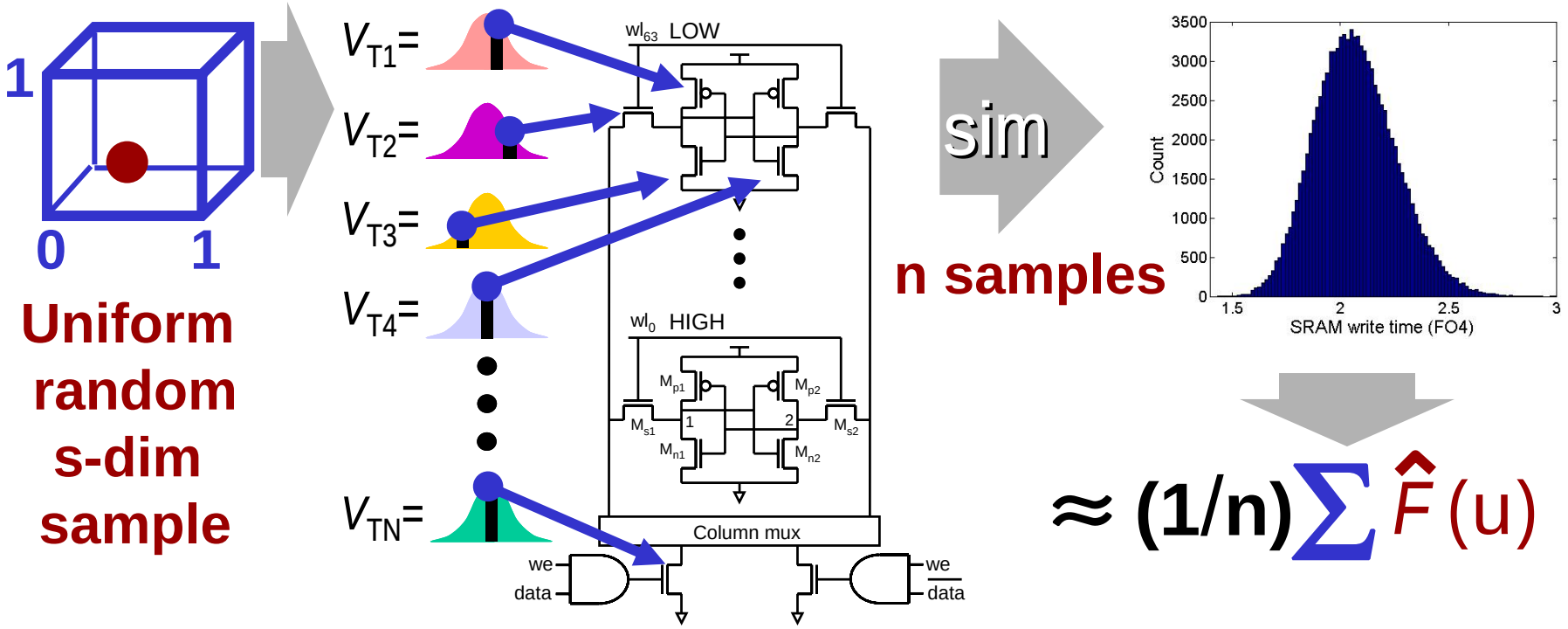
s-dimensional prob

$$\approx (1/n) \sum \hat{F}(u)$$



Can transform to sample uniformly from s-dim unit cube

Evaluate Circuit Impact: Monte Carlo



- "Unit cube" thing seems like a *minor* detail
- ... but it turns out to be *crucial*

Why is Monte Carlo Difficult?

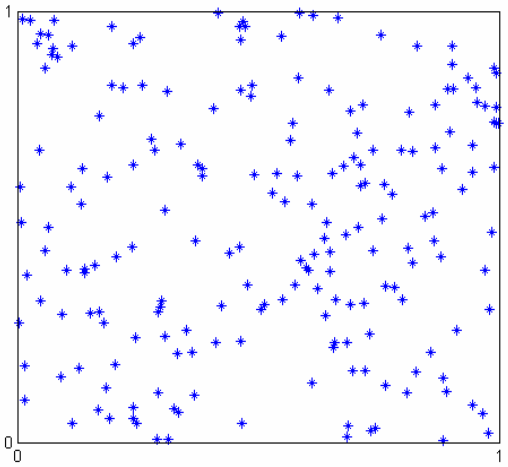
- **High-dim problems:** ***s is big (100-1000)***
- **Profoundly nonlinear:** ***Nanoscale physics***
- **Accuracy matters:** ***~1-5% error***
- **Speed matters:** ***Many samples***
- **Samples expensive:** ***Simulate each circuit***

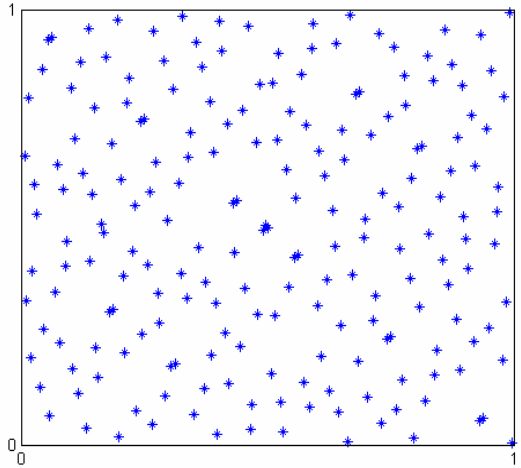
Question: Who *Else* Has This Problem?

Computational finance(!)

- **Valuing complex financial instruments, derivatives**
- **High-dimensional, nonlinear, statistical integrals**
- **Speed+accuracy matters here, e.g., ~real-time decision-making**

Big Idea: Quasi Monte Carlo (QMC)





- **Classical Monte Carlo**
  - **Uniform pseudo-random pts**
  - **Surprise: not very uniform**
- **Error for n samples**

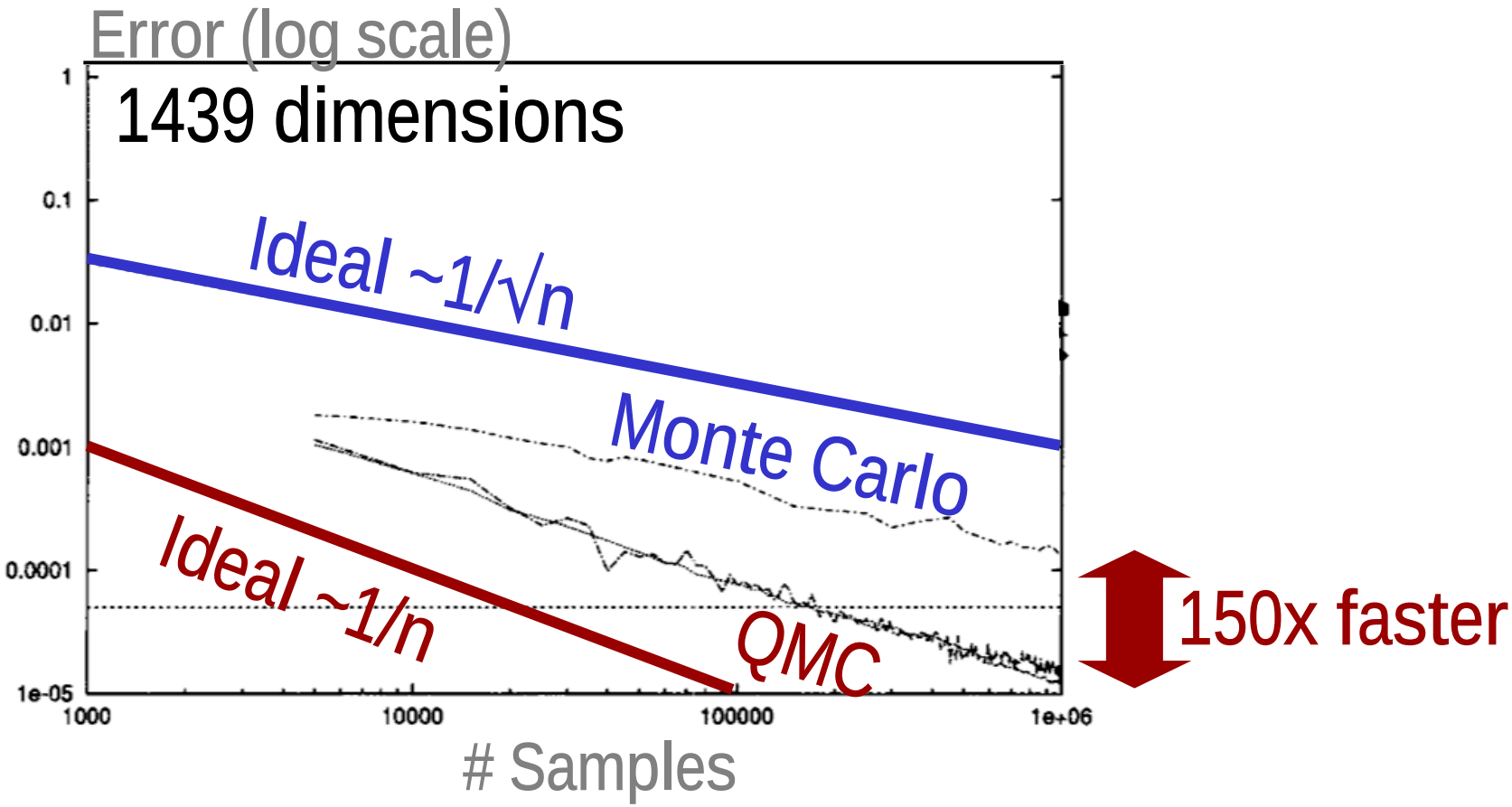
$$O(1 / \sqrt{n})$$

- **Quasi Monte Carlo**
  - **Deterministic samples**
  - **“Low-discrepancy” pts**
- **Error for n samples**

$$O(1 / n)$$

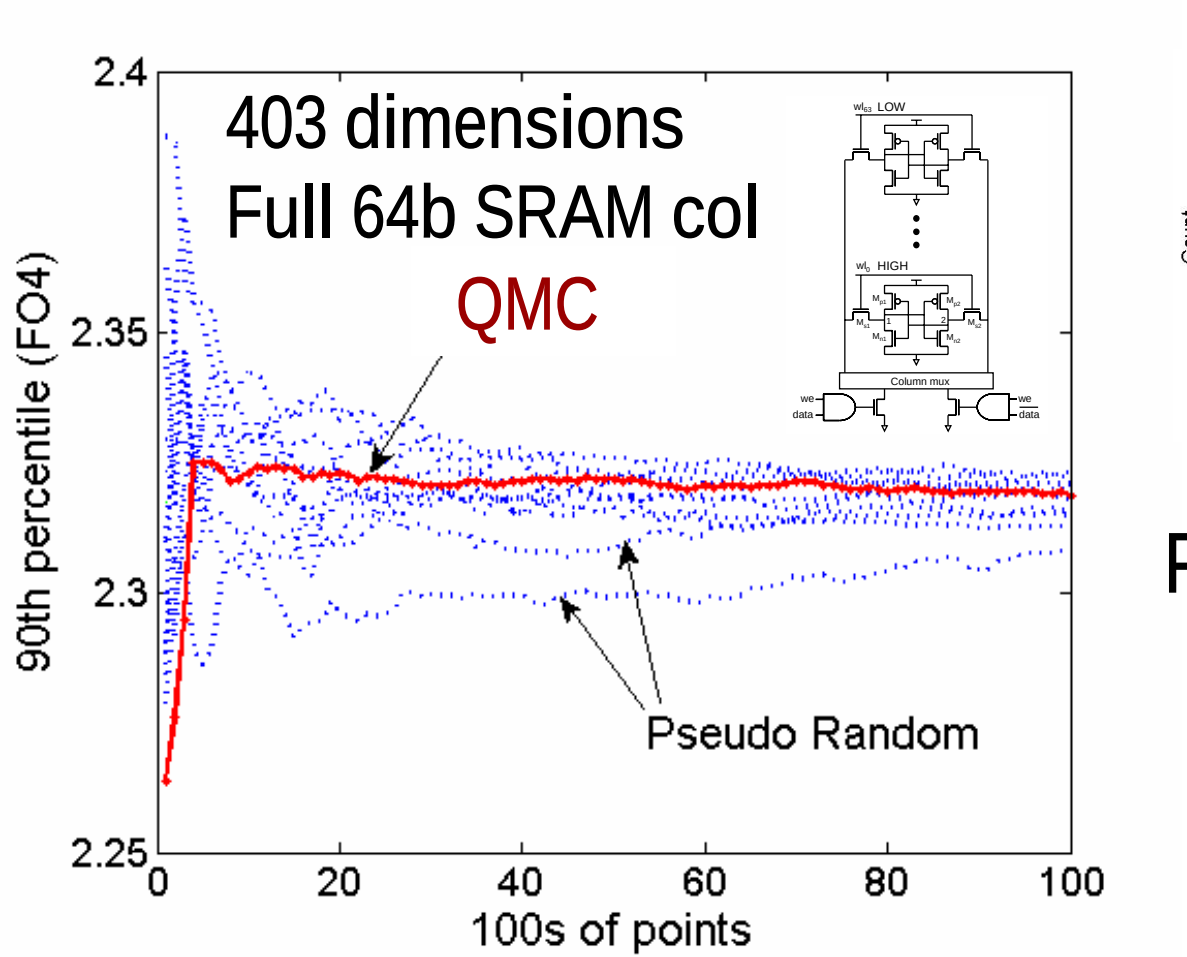
Computational Finance Example

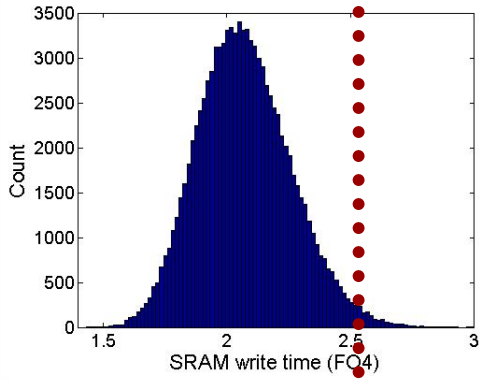
- **Eval 5-year discount price for a bond**
  - **From [Ninomiya,Tezuka, App Math Finance 1996]**



Does It Work for Circuits? (Yes!)

- **But requires some subtlety to map to QMC**
  - **[Singhee, Rutenbar, ISQED 2007, to appear]**

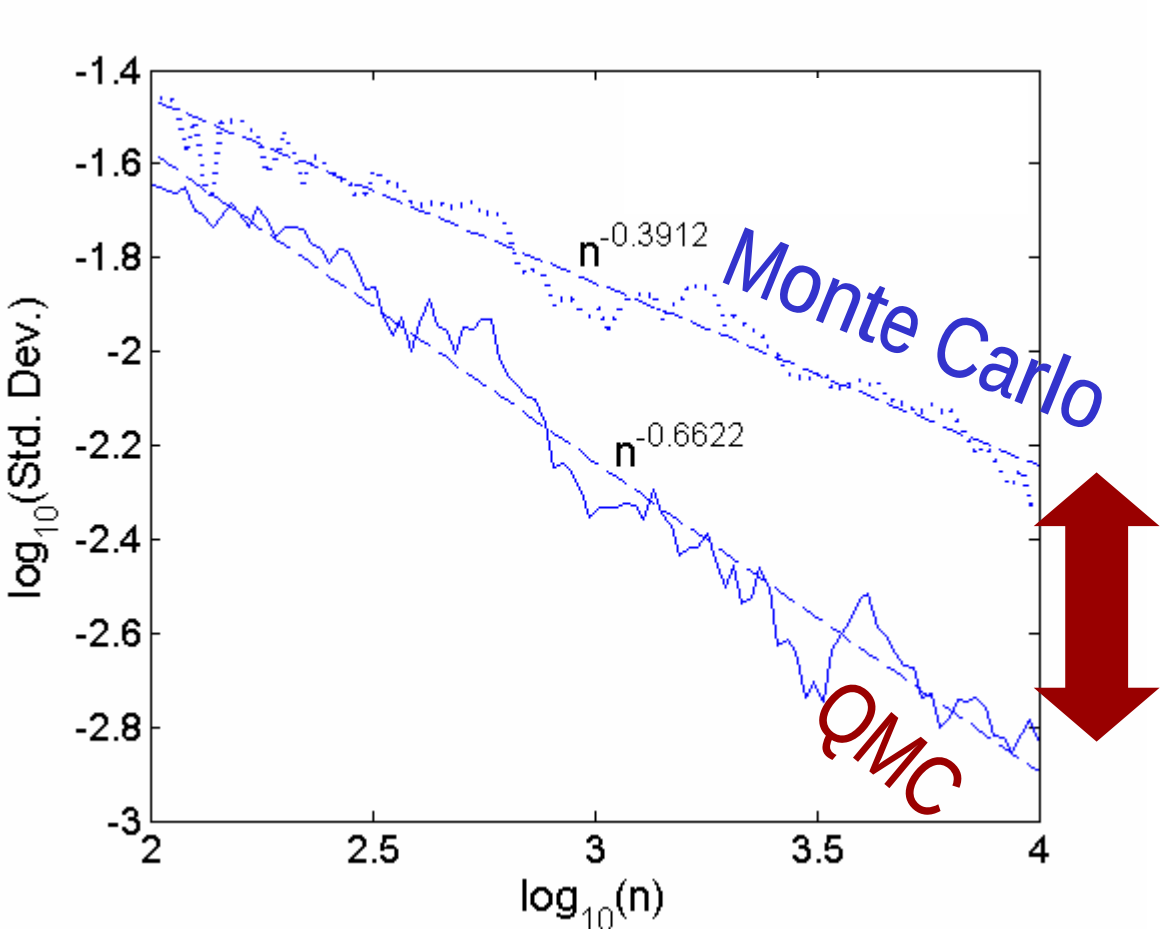




$Pr(write < t_w) = 0.9$

Very Promising Speedups

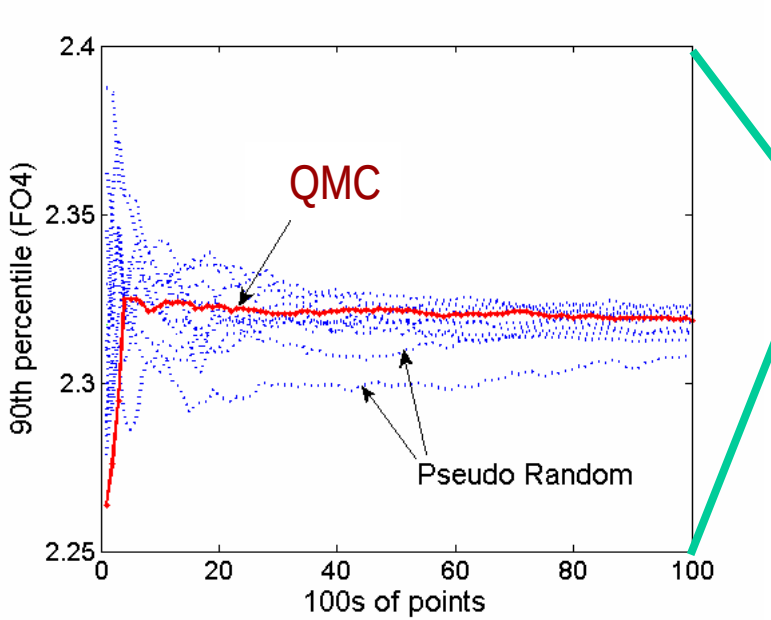
- **Same 403-dimensional, 64b SRAM column**

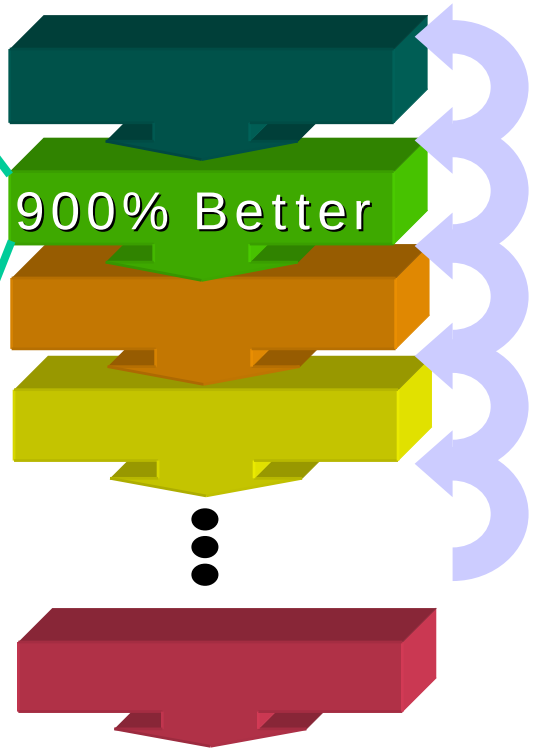


[Singhee, Rutenbar, ISQED 2007, to appear]

Rutenbar's Rule—Revisited

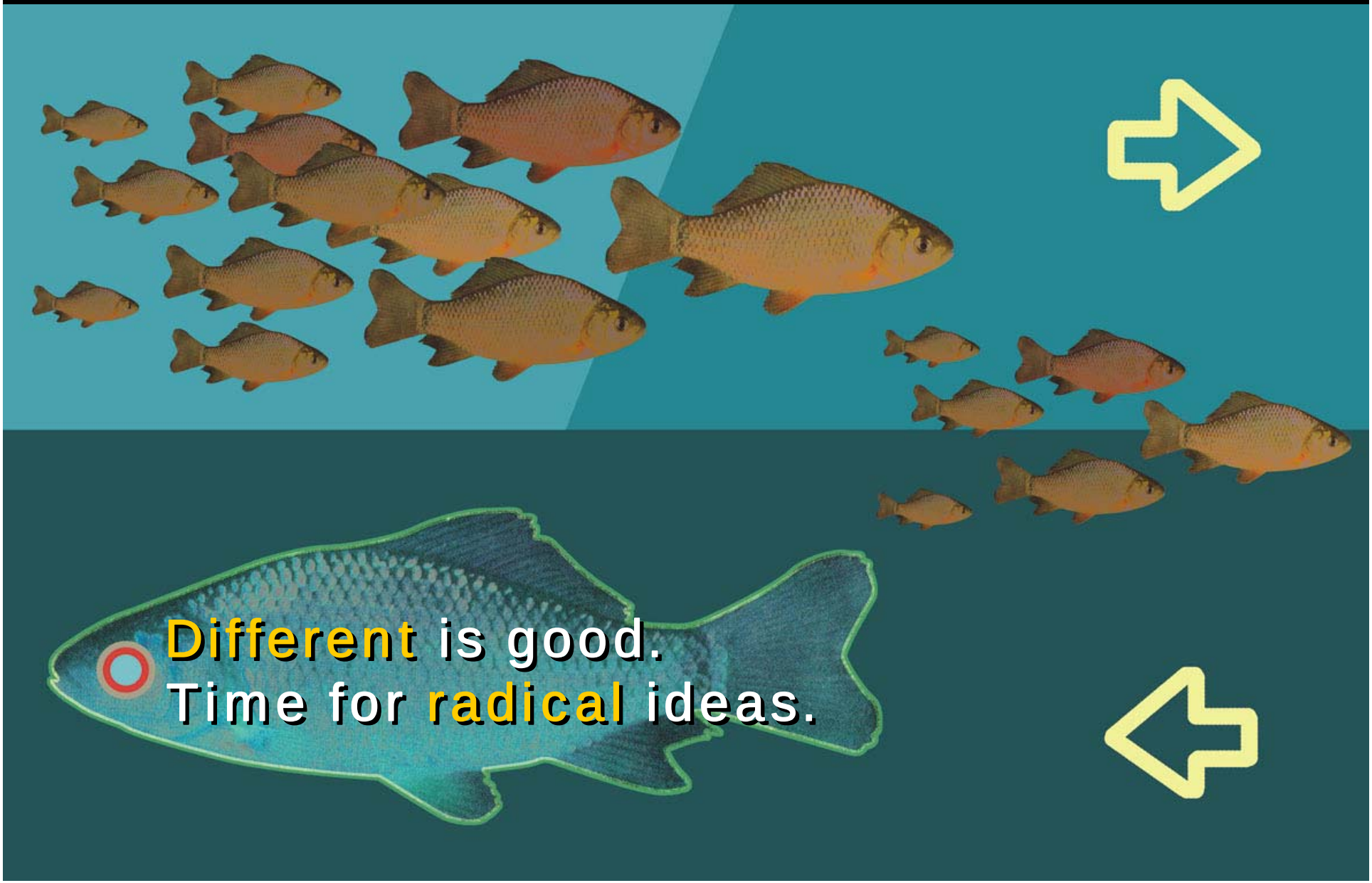
- **The taller the tool-chain, the more difficult for innovation in 1 tool to "survive the flow"**



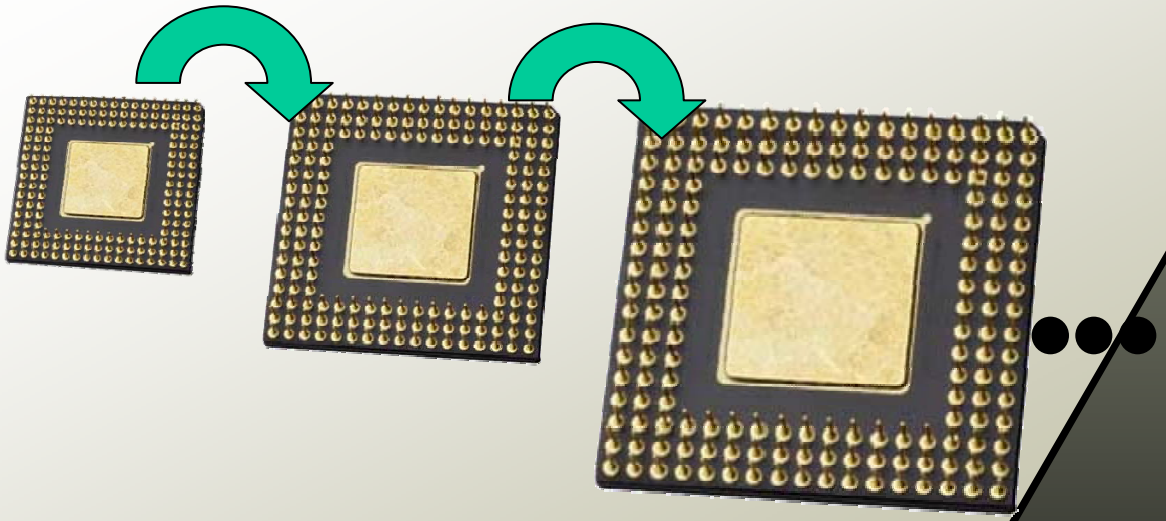


More Impact

“Tall Tool-Chains” Advice

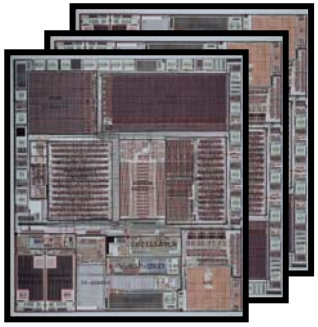


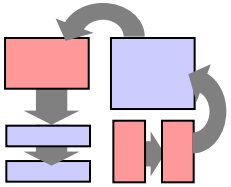
Big Challenge #3: Big Systems

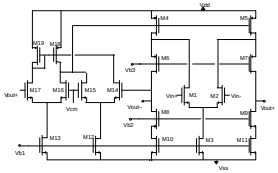


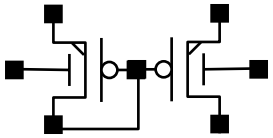
- Challenges in complexity

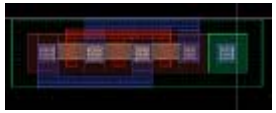
How To Explain (*Teach*) This...?

Systems  10M – 10B devices

Architectures  10K – 1M devices

Basic blocks  100-1000 devices

Circuits  10 devices

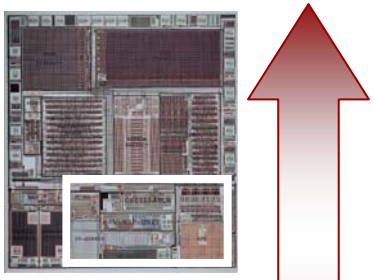
Devices  1 device

Unique problems as we go "UP"

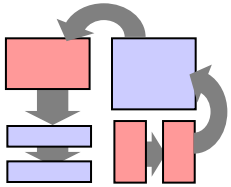
“Big Systems” Advice

- How do you learn to be a tightrope walker?
  - Buy a book? A Powerpoint talk? Buy the DVD?
- No: you just go do it, try it, practice it

Same with System Design: Just Do It

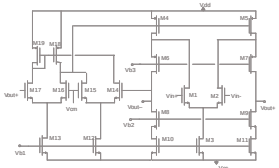


Systems

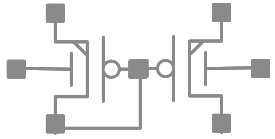


Architecture

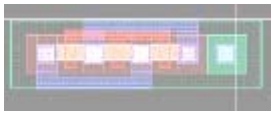
Go explore tradeoffs & consequences at top



Basic blocks



Fundamental circuits



Devices

Make some technology decisions at bottom

Form right team(!)

“Big Team” Example: Parallel Radios

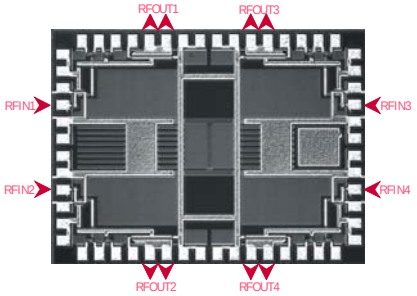
- **From Pat Gelsinger, Senior VP of Intel**

[Source: Pat Gelsinger, Intel]

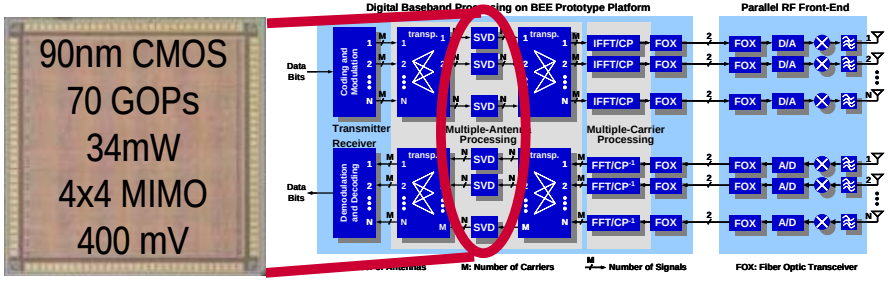
“Big Team” Ex: Parallel Radio Designs

- **Multi-university effort: Berkeley + MIT**
  - **US national FCRP Focus Center for Circuit & System Solutions (C2S2), funded by govt + semiconductor industry**

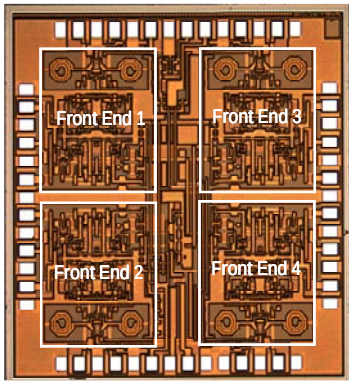
Parallel SiGe power amps



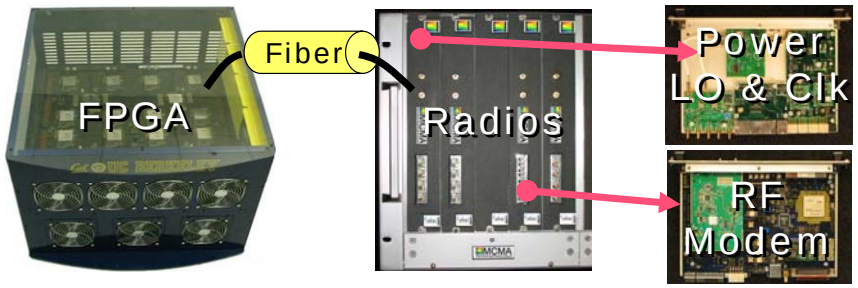
Parallel CMOS radio baseband DSP



Parallel SiGe RF frontends

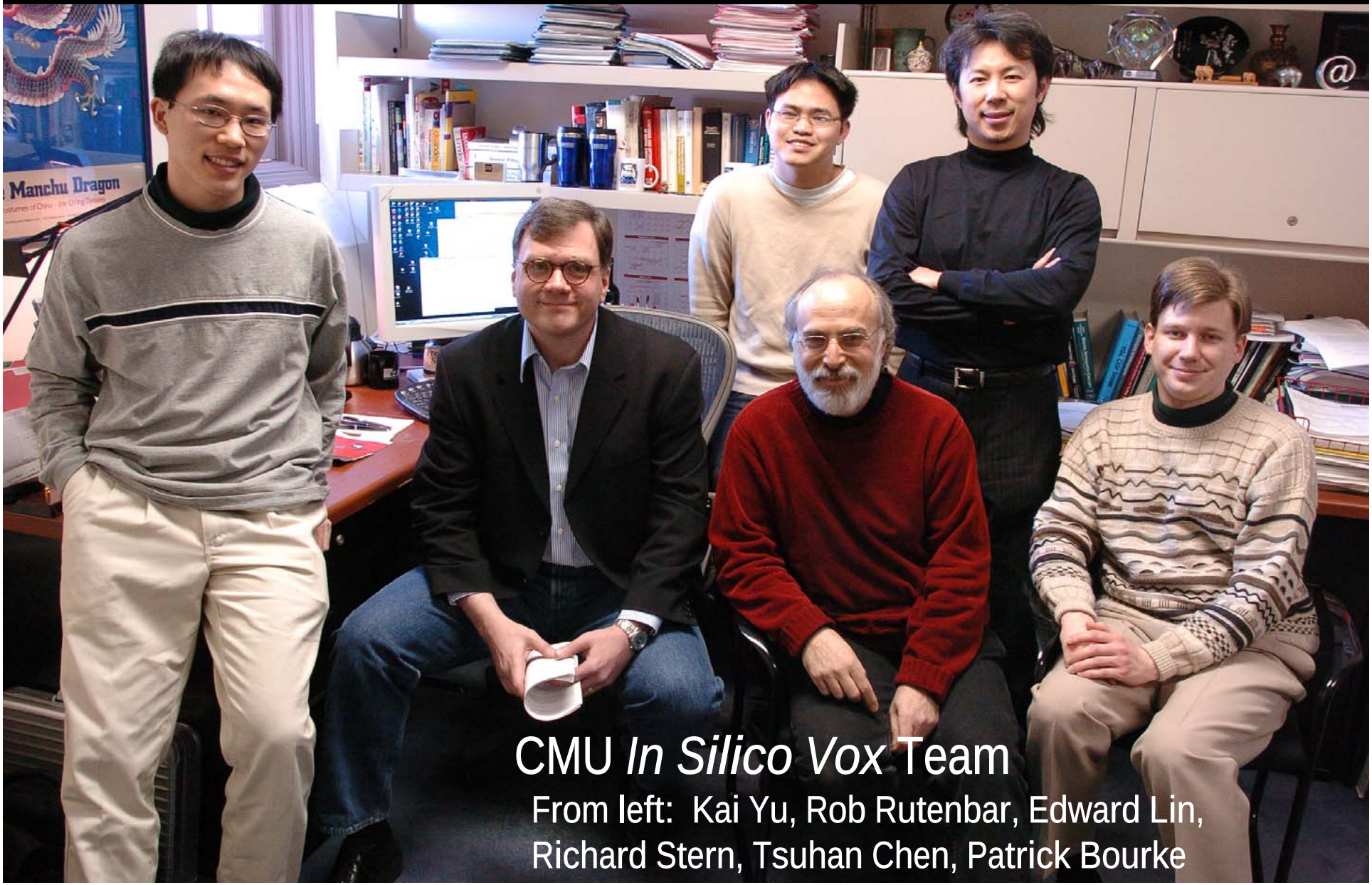


Parallel Radio Emulation System: BEE



[Charles Sodini, Greg Wornell, MIT] [Bob Brodersen, Bora Nikolić UCB]

But A Small Team Can Do It Too...



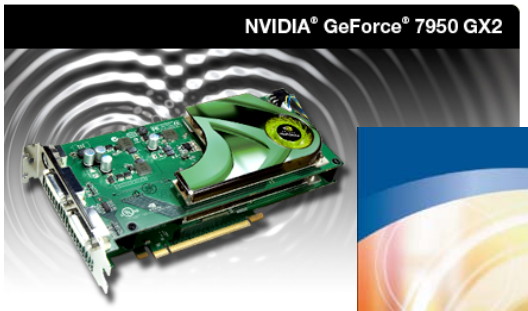
CMU *In Silico Vox* Team

From left: Kai Yu, Rob Rutenbar, Edward Lin, Richard Stern, Tsuhan Chen, Patrick Bourke

In Silico Vox: Speech Recognition in Silicon

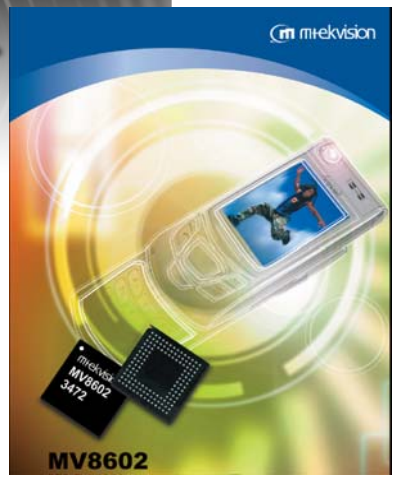
- **Paint pixels in software?**
- **No! Graphics chips**

True on the desktop



http://www.nvidia.com

& cell



http://www.mtekvision.com

- **So why are all today's best speech recognizers done in *software*?**

Can we do better?

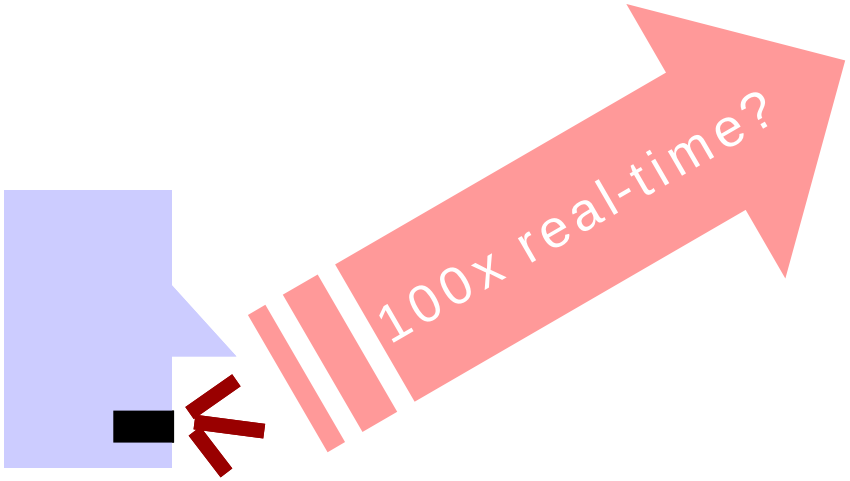
Next-Gen Apps Need >100x Improvement

Audio-mining

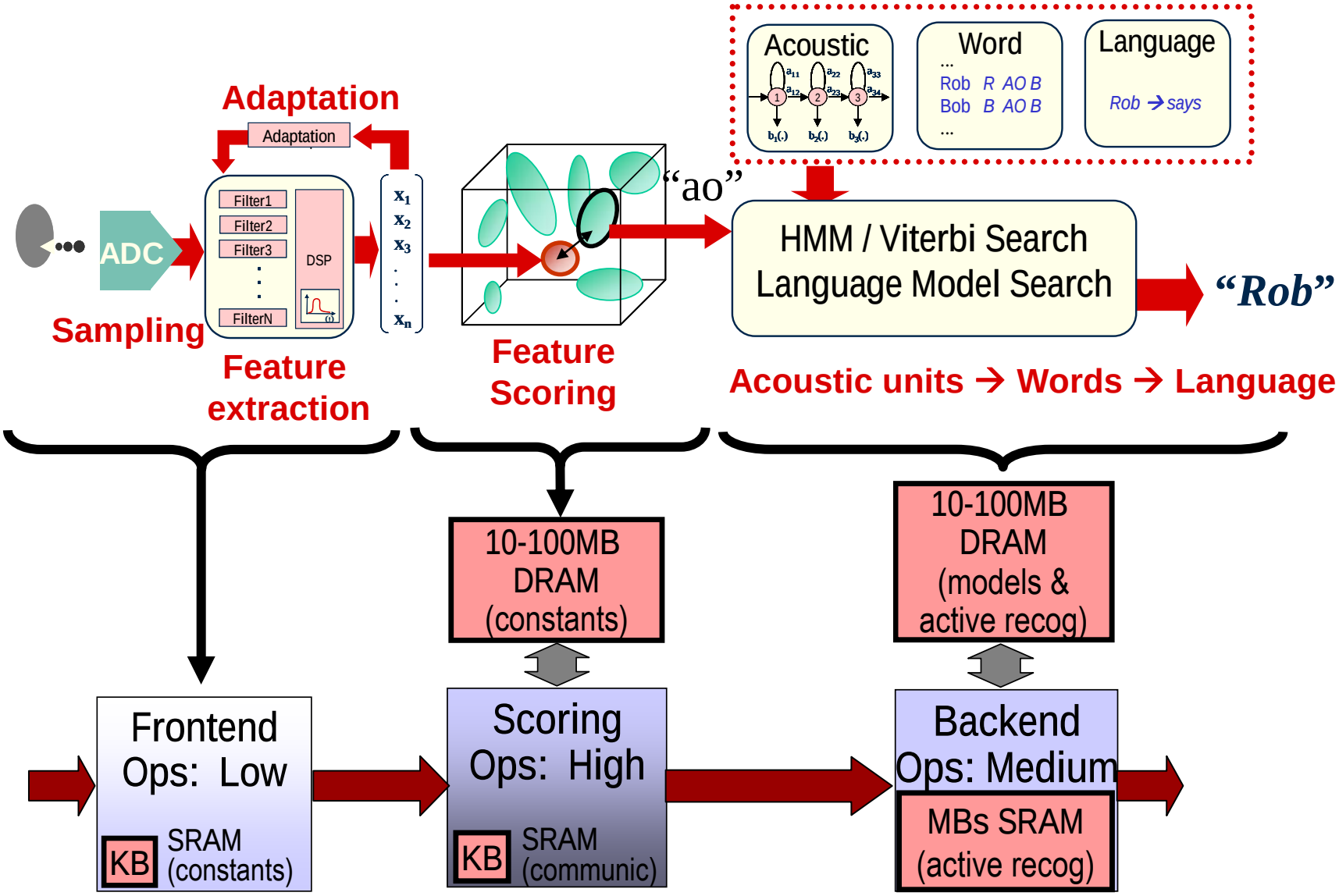
- **Very fast recognizers –faster than realtime**
- **App: search media streams (DVD) quickly**

Hands-free appliances

- **Very portable recognizers –high quality on << 1 watt**
- **App: interfaces to small devices, cellphones**

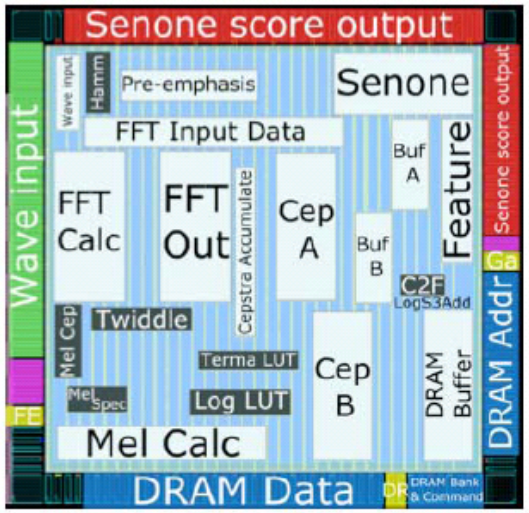


Speech: Complex Task to do in Silicon

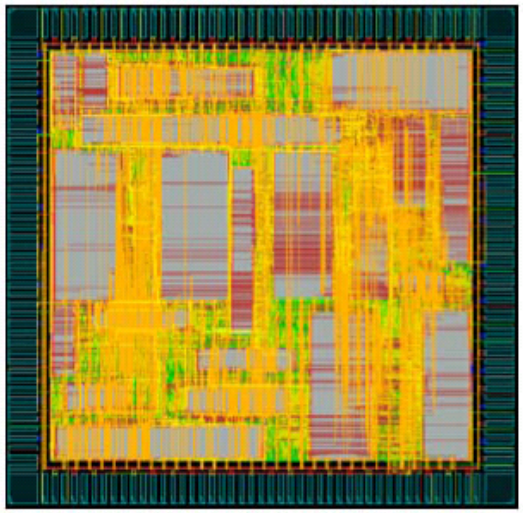


Pieces of Design: Great Class Projects

- **CMU student team: Patrick Chiu, David Fu, Mark McCartney, Ajay Panagariya, Chris Thomas**



Floorplan



Final Layout

Area	11.16 mm^2 core / 16.09 mm^2 chip
Effective Utilization	53.32%
Cell Rows	657
Cells	67354
Pins	225358
IO Pins	94
Nets	79382
Avg. Pins/Net	2.84
Nets	
(Internal)	77977
(External)	94
Connections	
(Internal)	146621
(External)	188
Total net length	6.00 m
(X)	2.59 m
(Y)	3.40 m
Power Supply	1.98 V
Average Power	19.8 mW
(switching)	11.78 mW
(internal)	7.98 mW
(leakage)	0.036 mw
Power by clock domain	
Frontend	2.018 mW
Gaussian	14.25 mW
DRAM	2.57 mW
Unclocked	0.96 mW
Power by cell category	
Core	19.5 mW
Block	0.29 mW
IO	0 mW
Worst IR drop	0.012 V

Final Stats

CMU *In Silico Vox* Project: FPGA Results

- ***Most complex*** **recognizer ever mapped to hardware**
  - **[Lin, Yu, Rutenbar, Chen, HOTCHIPS 2006]**

“Big Systems” Advice: Tightrope Walking

- **How to learn to deal with system complexity?**
- Pick **a complex system.** Deal **with it.**

Concluding Thought

- "One thing we know about **creativity** is that it typically occurs when people who have mastered two or more quite **different** fields use the framework in one to think afresh about the other...

  Marc Tucker, *Tough Choices or Tough Times*

- Lithography + EDA optimization
  ... Semiconductors + Computational finance
  ... Speech recognition + Custom silicon
  ... Many creative combos await discovery

Thank You!

Acknowledgements

- **My graduate students, for contribs to talk**
  - **Edward C. Lin, Amith Singhee, Sonia Singhal, Kai Yu**
- **My many faculty and student colleagues in C2S2 Focus Center (www.c2s2.org)**
- **Funding from FCRP, SRC, NSF, DARPA, and US semiconductor industry**