Variability-Aware Product Development in a Fabless Environment

Jeff Watt

May 13, 2009

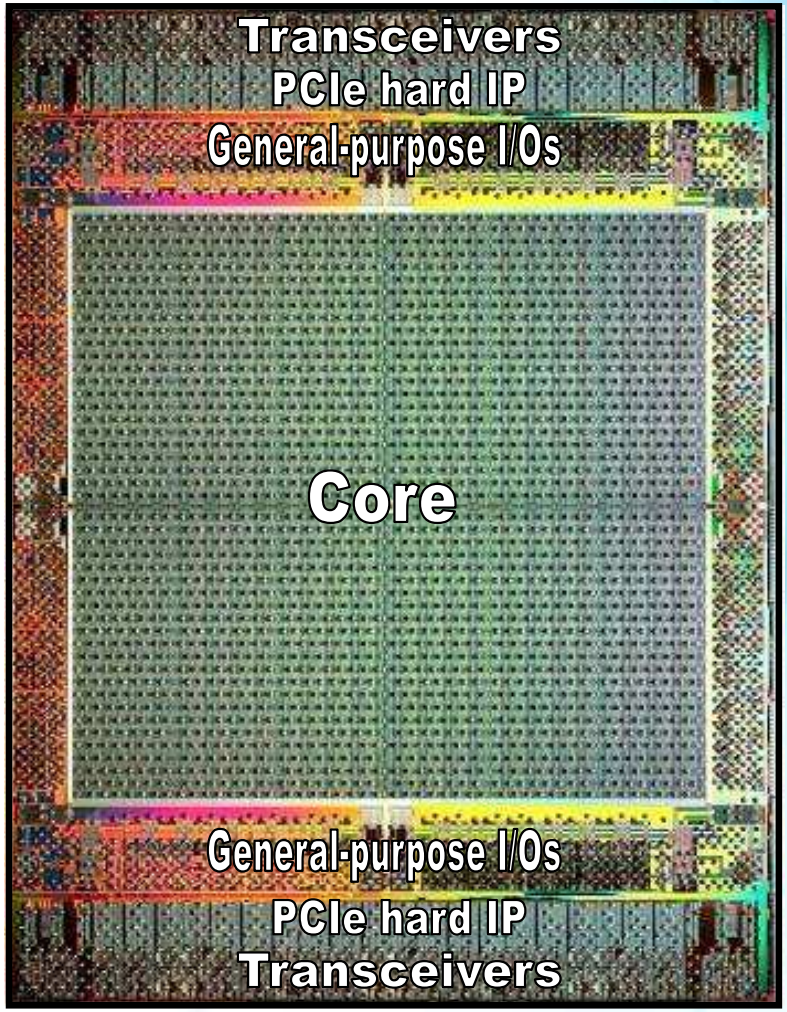
Outline

- Scaling challenges
- Collaborative development approach
- Variation modeling methodology
- Test-chip program
- Variability outlook
- Summary

Leading-Edge FPGA Attributes

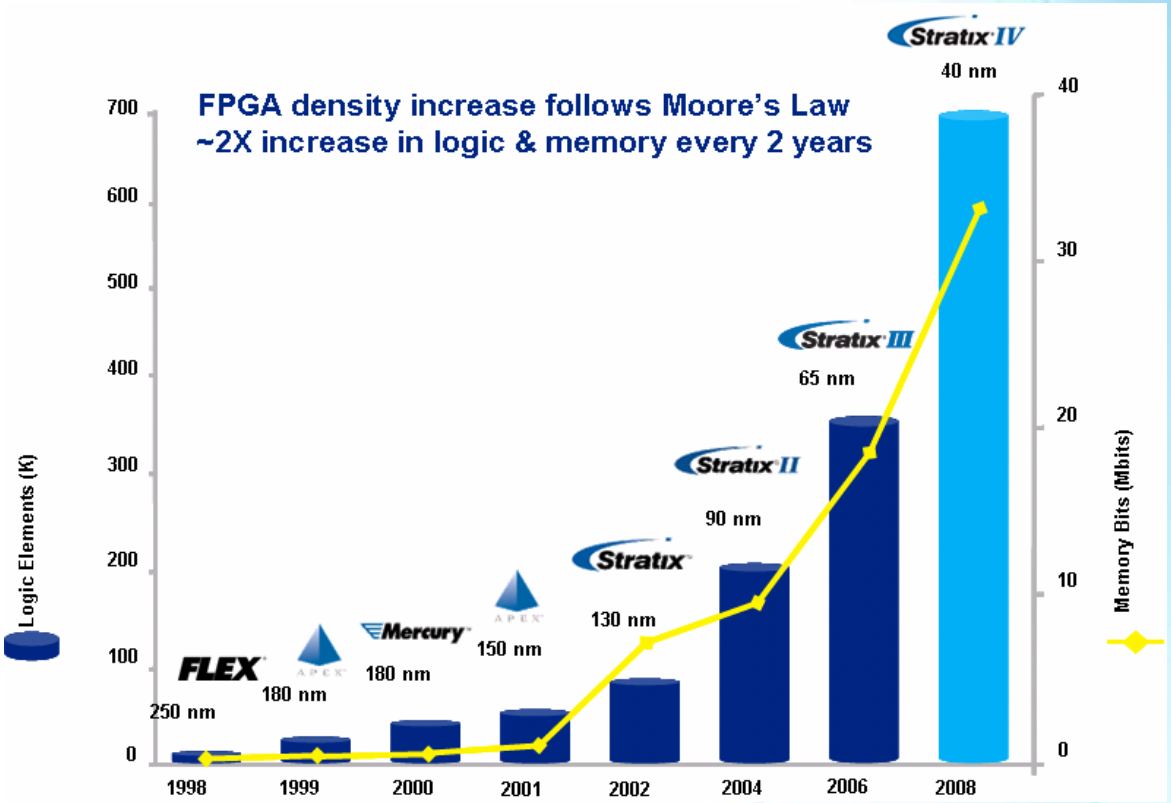
- Advanced 40-nm foundry process:
  - Triple gate-oxide
  - Multiple threshold voltages
  - 2nd-generation strained silicon
  - 12-level metal
- Very high density
  - Up to 2,500,000,000 total transistors
  - Up to 32 Mbits of user dual-port SRAM
  - Up to 100 Mbits of configuration memory
- Primarily full custom design
  - Limited use of foundry standard cells
- Low core supply voltage: 0.9V
- Aggressive development schedule
  - First product shipped to customers Dec. 2008

Stratix IV GX 40-nm FPGA



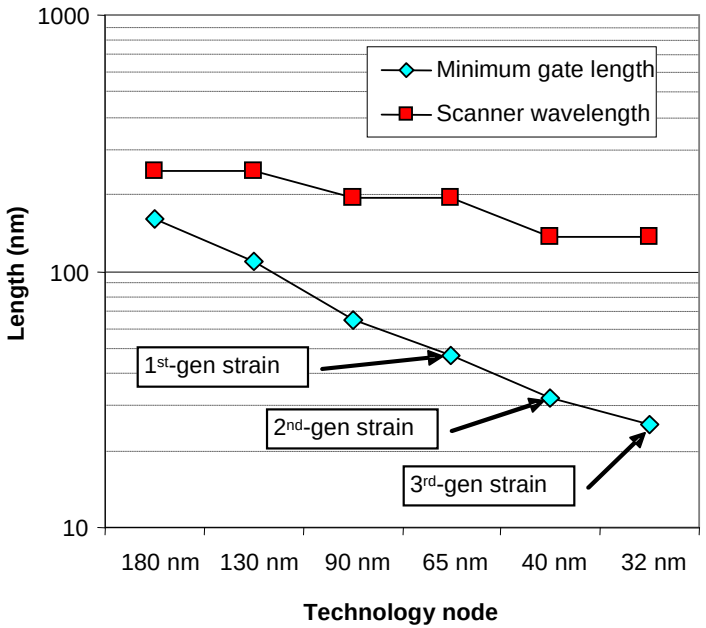
Scaling: The Good, the Bad, and the Ugly

- Good:
  - Higher density
  - Higher performance and throughput
  - More features and functionality
  - Lower power per function
  - Lower cost per function
- Bad and Ugly:
  - Increasing variability
  - Increasing complexity
  - Increasing cost (development and wafer)
  - Higher total power (for fixed die size)

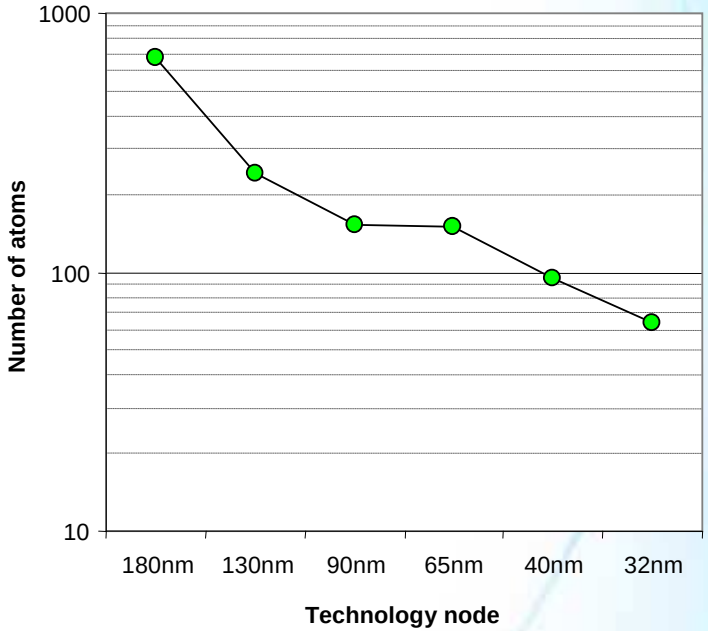


Scaling Exacerbates Within-Die Variation

Minimum gate length and lithography wavelength

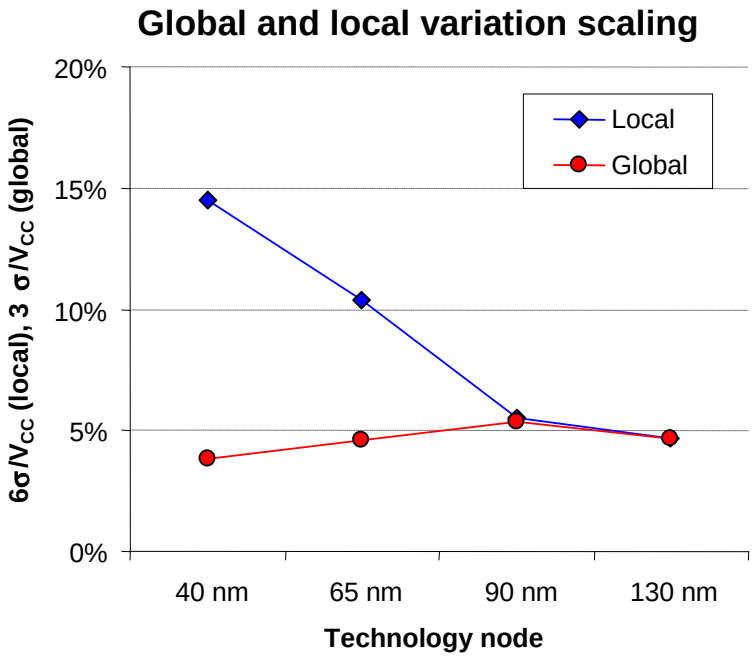


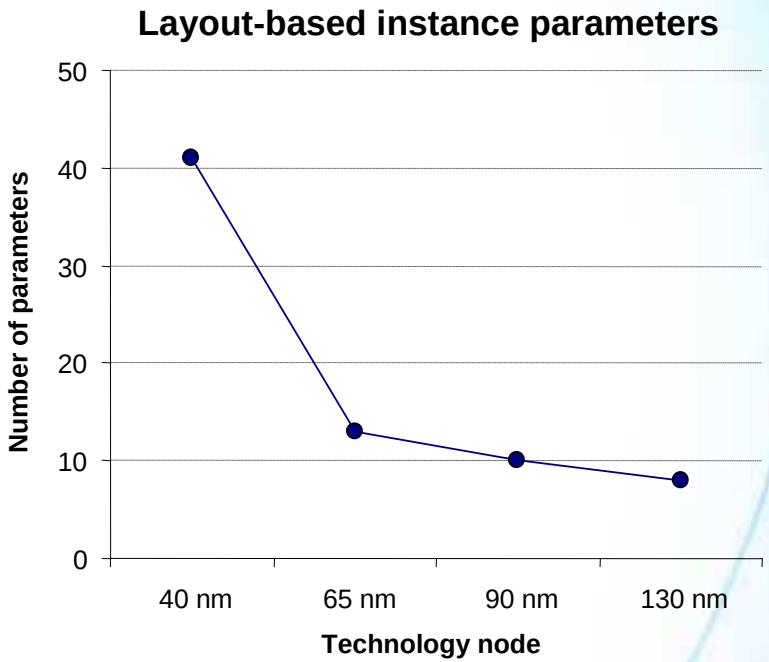
Number of dopant atoms in minimum geometry channel



- Scaling trends enhance within die variation
  - Widening gap between gate length and lithography wavelength degrades pattern fidelity and increases layout sensitivity
  - Process-induced strain enhances layout sensitivity
  - Reduction of number of dopant atoms in channel increases local variation

Variability and Complexity Trends





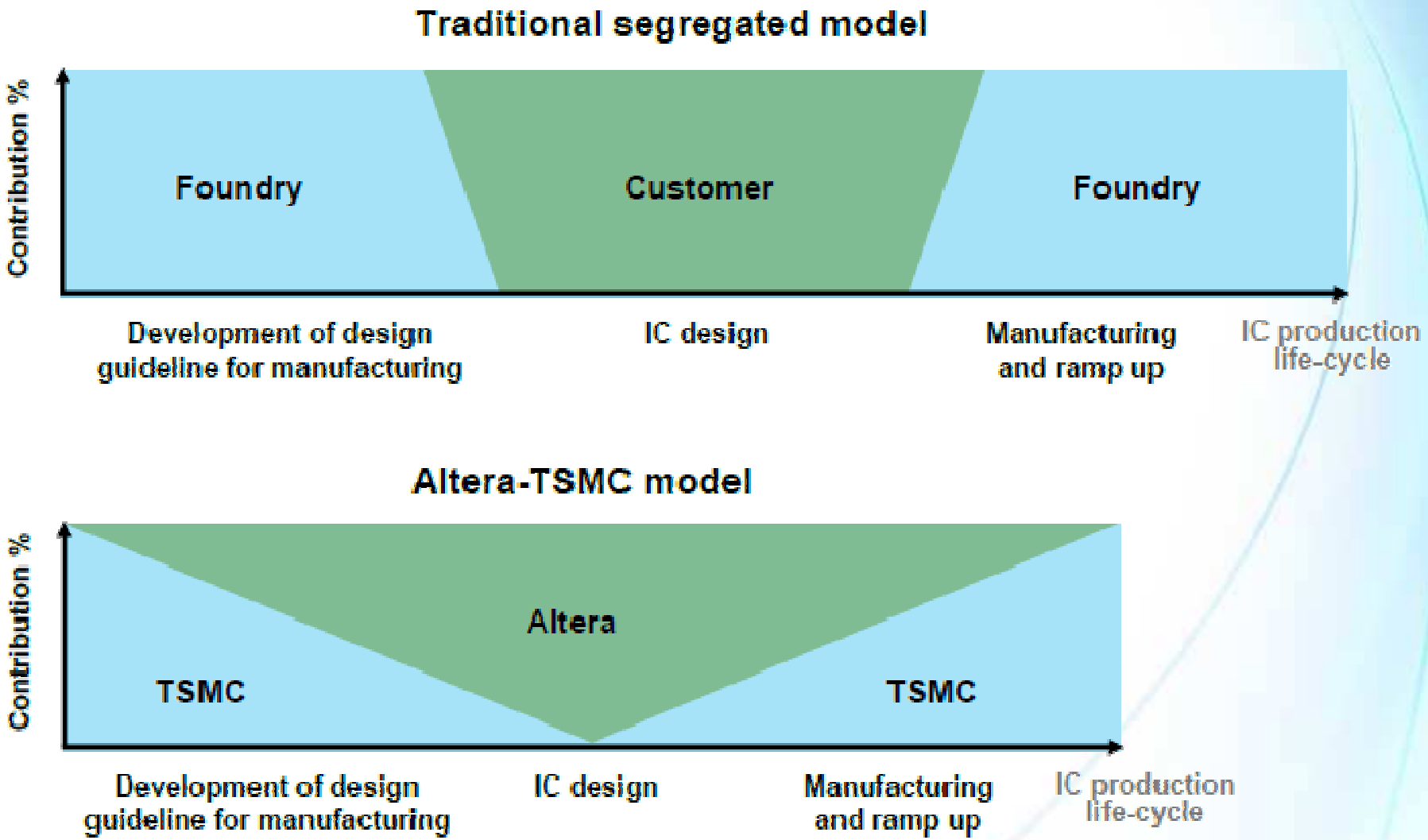
- Local variation increases with technology scaling due to reduced transistor channel area
  - Sensitivity to local variation also increases due to reducing supply voltage
- Global variation stays constant due to improvements in process control despite smaller geometries
- Model complexity increases due to layout dependence

Fabless IC Development Challenges

- Process-design interactions are increasing
  - Layout-induced variability and increased sensitivity to process variation
- But access to process information is limited
  - Foundry must protect proprietary process details
- Early adopters develop product while process is immature
  - Volume statistical data is unavailable
  - Process will evolve during development
- But “first silicon to production” approach is needed
  - Driven by competitive pressure and cost of re-spin

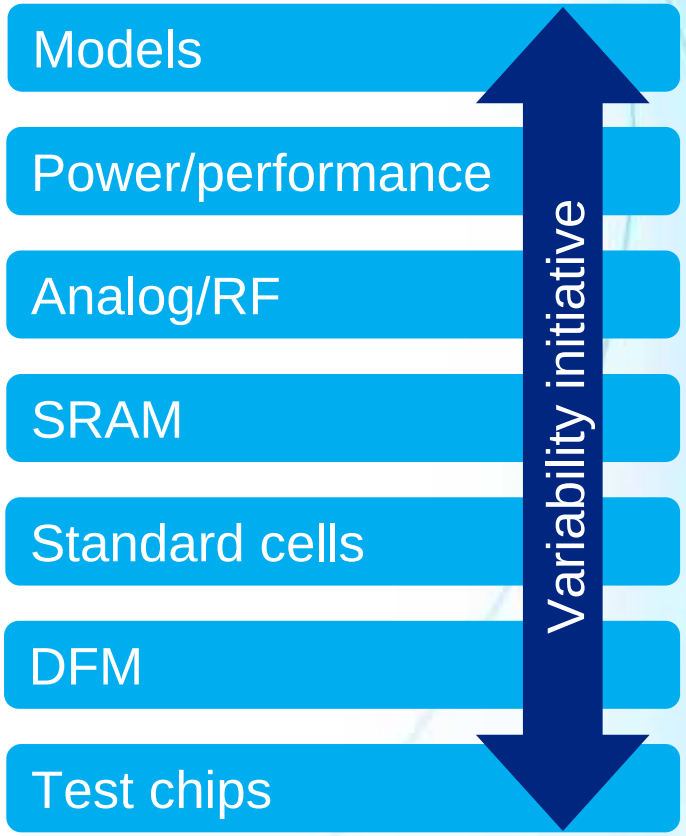
Solution:
Comprehensive variability-aware collaborative development

Collaborative Development Model



“Tiger” Team Methodology

- Variability initiative goals
  - Accurate models
  - Robust simulation methodologies
  - Identify and focus on highly sensitive components such as embedded SRAM
  - Test-chip validation
- Tiger teams address all critical aspects of development
  - Each team has a specific goal, Altera team leader, and TSMC team leader
- Variability is a cross-cut initiative spanning several teams



Transistor Variation Components

Component	Type	Source	Modeling approach
Layout-dependent	Systematic within-die	Pattern density and fidelity, Dopant non-uniformity, Process-induced strain	Extract from layout
Local	Random within-die	Random dopant fluctuations	Statistical model
Global	Random die-to-die	Spatial and temporal process variations	Worst-case corners or statistical model

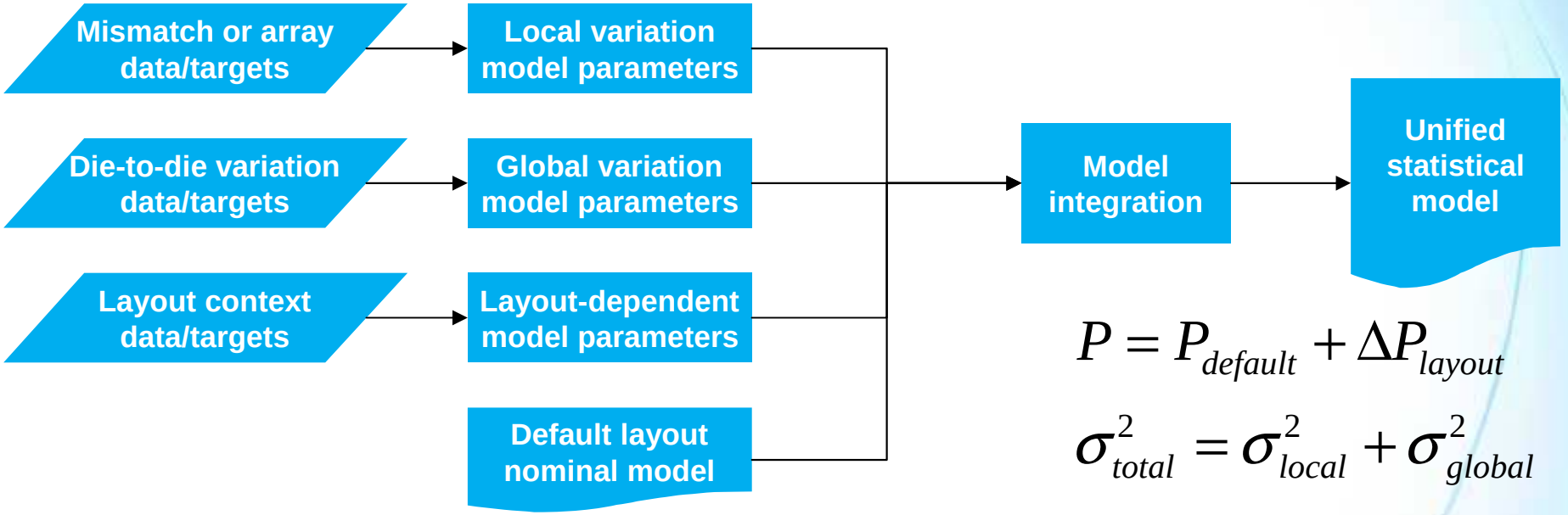
- Variation modeling philosophy:
  - Incorporate all systematic components into model and extract parameters from layout
  - Implement robust statistical simulation methodology for random components

Statistical Modeling Progression

Technology node	Statistical modeling approach	Applications
90 nm	Mismatch	Analog circuits
65 nm	Independent local and global statistical models	Analog circuits Memory blocks
40 nm	Unified statistical model	Analog circuits Memory blocks Logic

- Enhancements in statistical modeling capability are essential as technology scales
  - Design margins are decreasing
  - Local variation is increasing
- Unified statistical model approach introduced on 40 nm to address need for more accurate simulation of variation

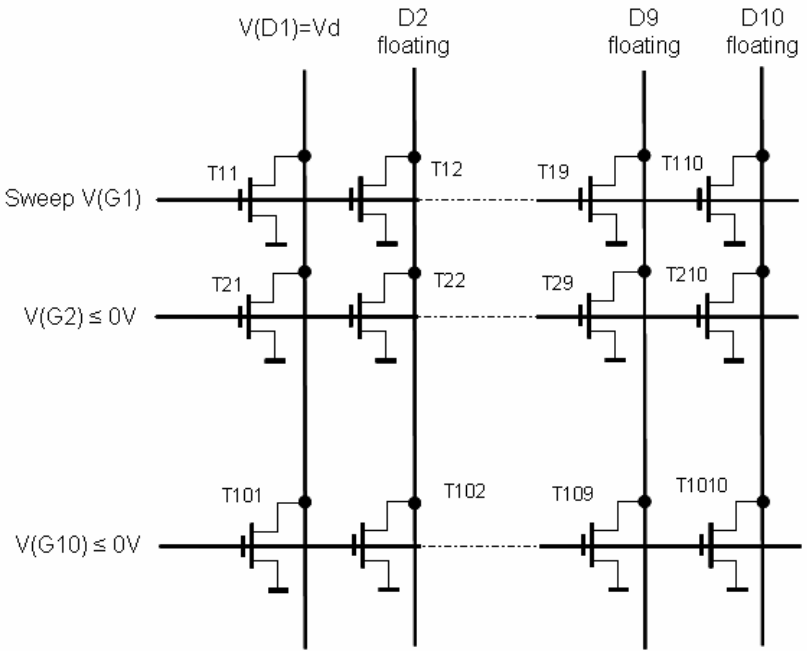
Unified Statistical Model



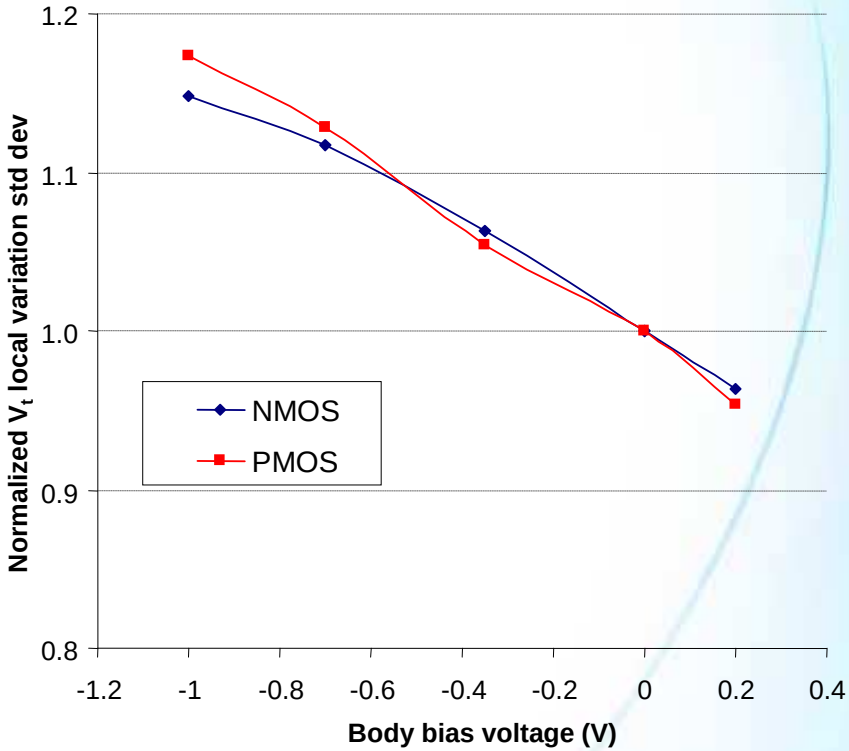
- Unified statistical approach incorporates all components of variation in a single model deck
- Modeled variation sources:
  - Random variation
    - Global (die-to-die): within wafer, wafer-to-wafer, lot-to-lot, fab-to-fab
    - Local (within die): variation obtained from mismatch or array structures
  - Systematic variation
    - Process induced: treated as part of global random variation
    - Layout induced: center shift (no added variation) based on layout extraction

Local Variation Characterization

10x10 directly addressable FET array



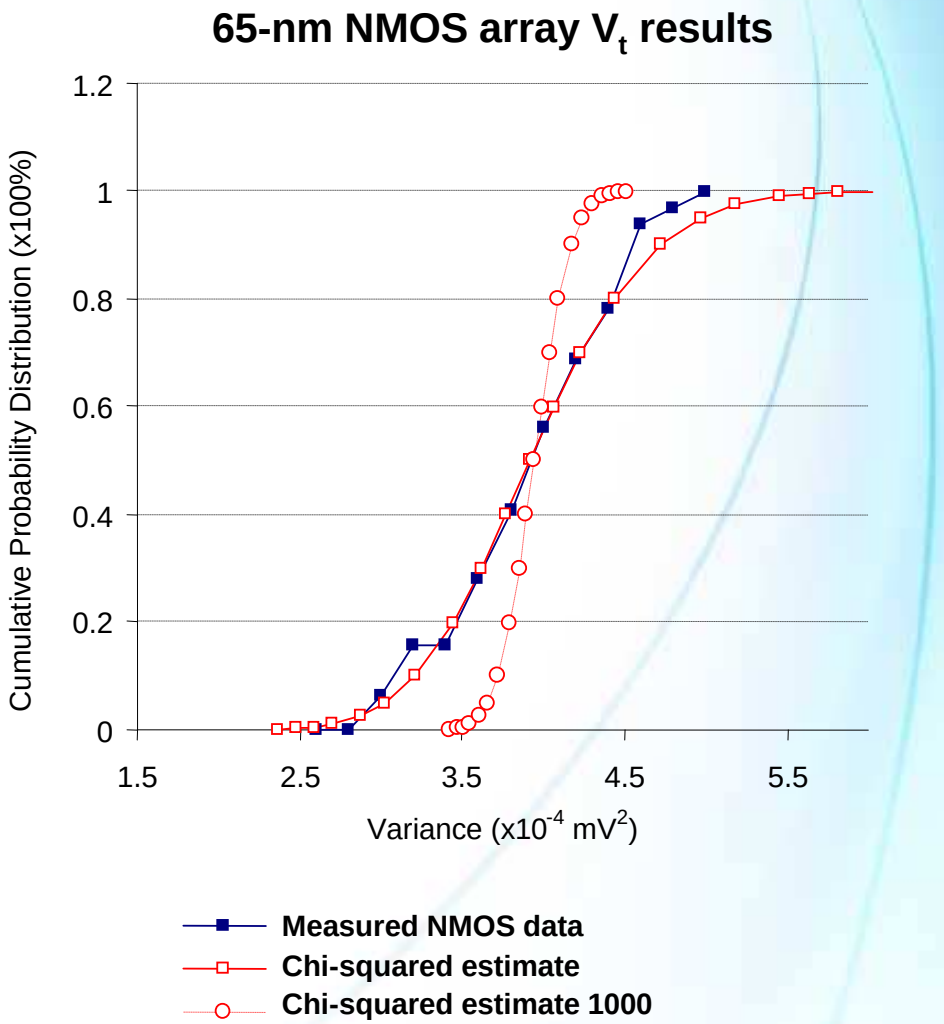
65-nm local variation sensitivity to body bias



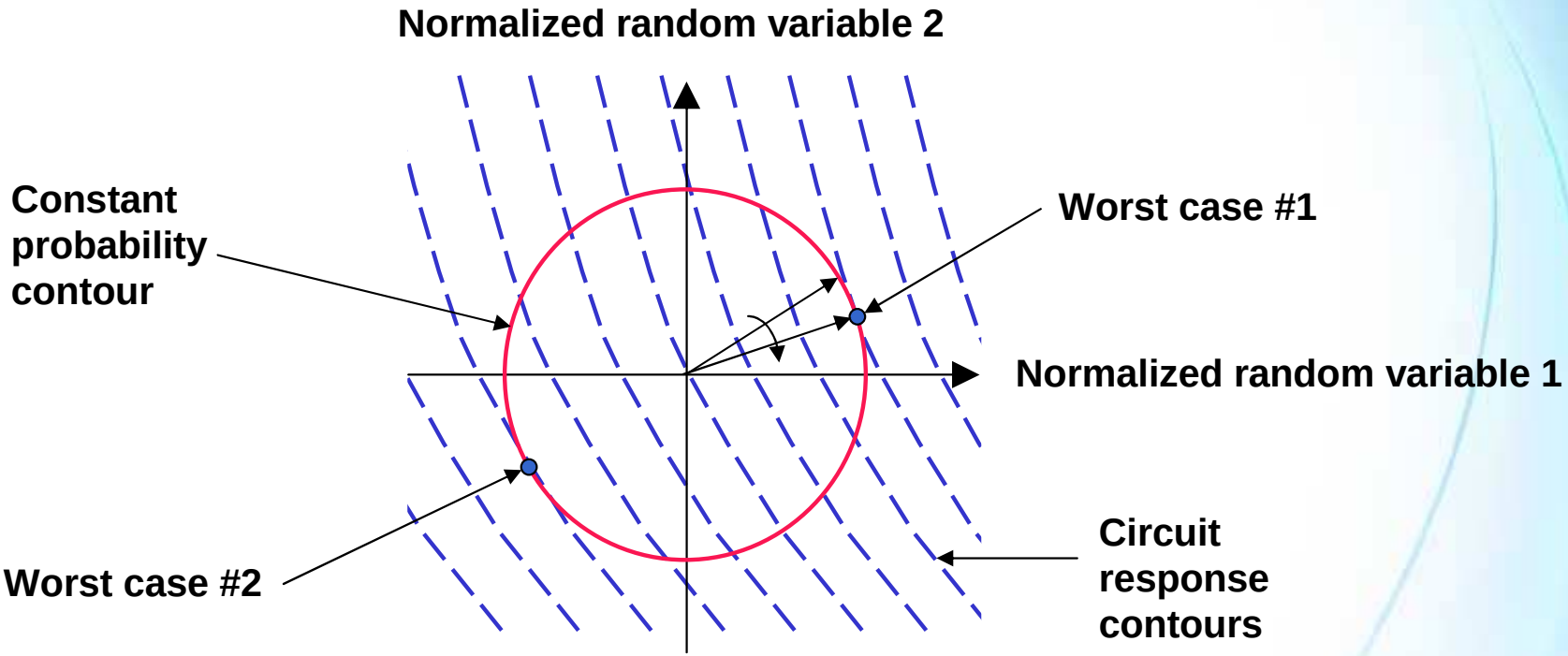
- Characterization supplements foundry data to target unique requirements of fabless product
  - Example: ensure accuracy across body bias used for programmable power control

Uncertainty of Variation

- Care must be taken when extracting variation model parameters
- Sample variance is only an estimate of true population variance
  - Sample variance distribution is defined by chi-squared distribution
- Sample size must be based on required confidence interval, or guardband should be applied

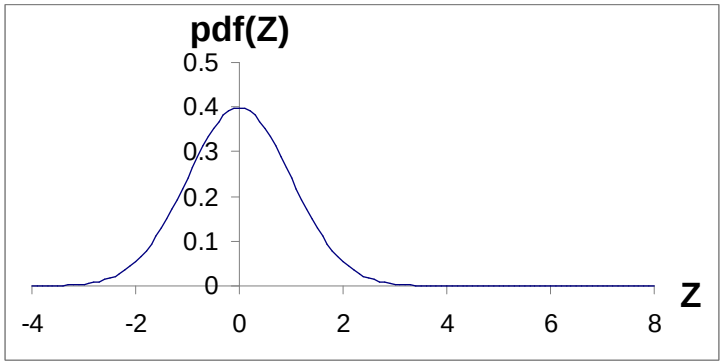


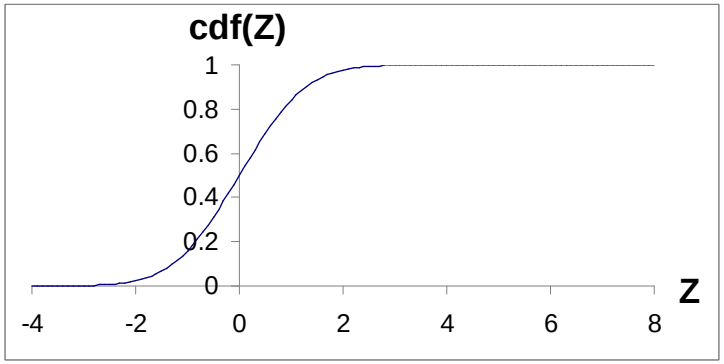
Statistical Simulation Methodology

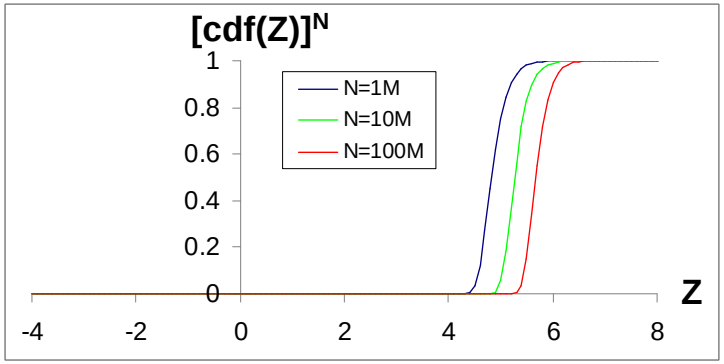


- Iterative method to locate the combination of principal components at a specified probability corresponding to worst-case circuit response
- In 2-dimensions: 2 points where the circuit response contours lines are tangent to the constant probability circle
- In N-dimensions: 2 points where the circuit response contour hyperplane is tangent to the constant probability hypersphere

How Many σs Are Enough?

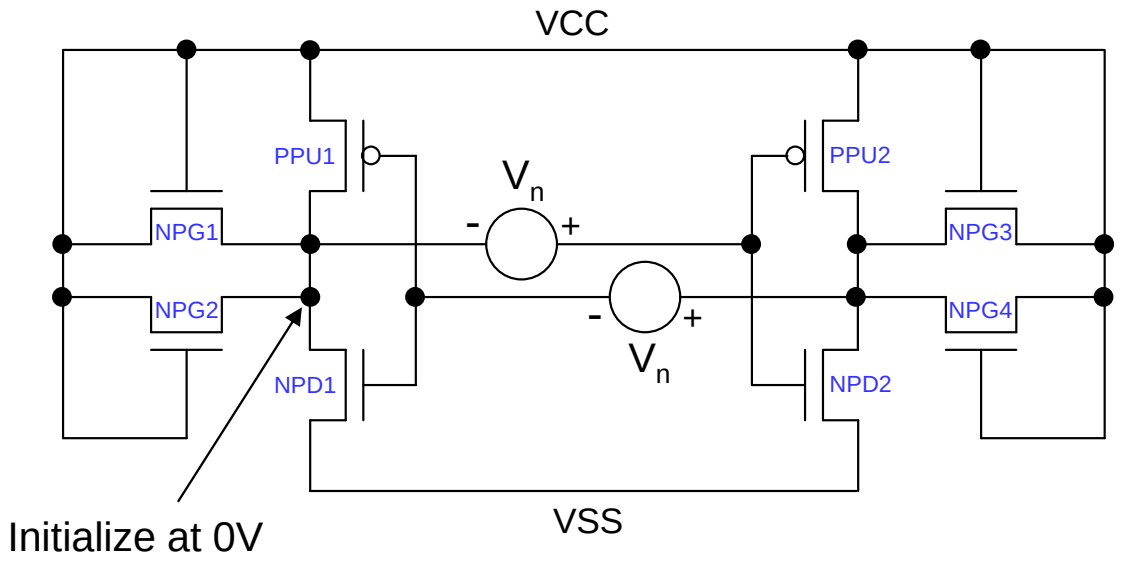






- 99% yield for one component per die
  - 2.3σ
- 99% yield for multiple components per die
  - 1M: 5.6σ
  - 10M: 6.0σ
  - 100M: 6.4σ
- Applies to any identical components which must simultaneously work (e.g., SRAM)
- Design tools enable product development team to use appropriate σ count to meet yield goals
- Design team education required to change mindset from 3σ to >6σ approach

65-nm Dual-Port SRAM SNM Example

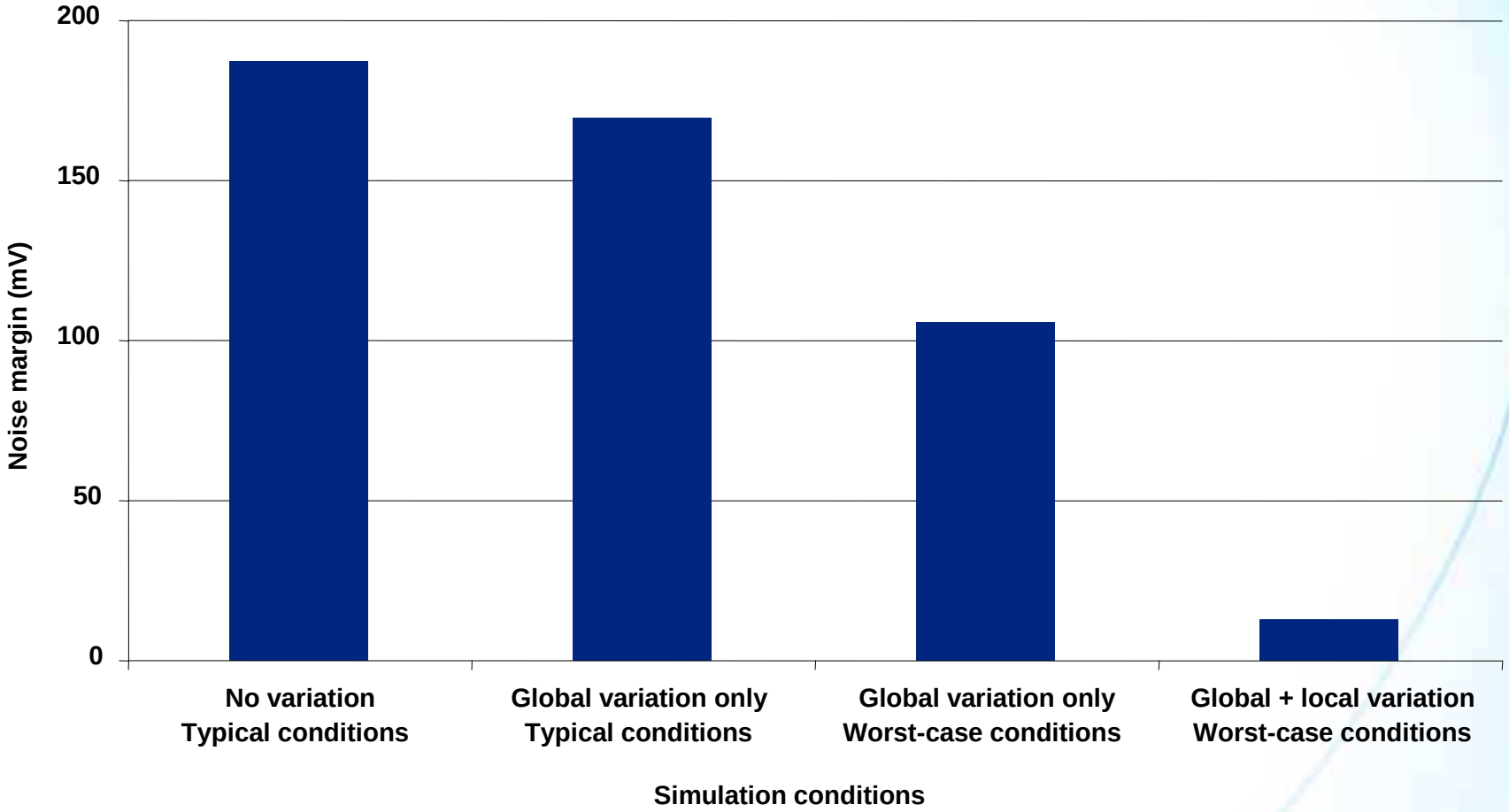


Static noise margin (SNM) = maximum V_n for stability

Test care	Worst-case cell transistor Vt shift (# of standard deviations)							
	NPD1	NPD2	NPG1	NPG2	NPG3	NPG4	PPU1	PPU2
SNM	3.49	-3.84	-1.86	-1.86	0	0	0	-1.47

6σ local variation

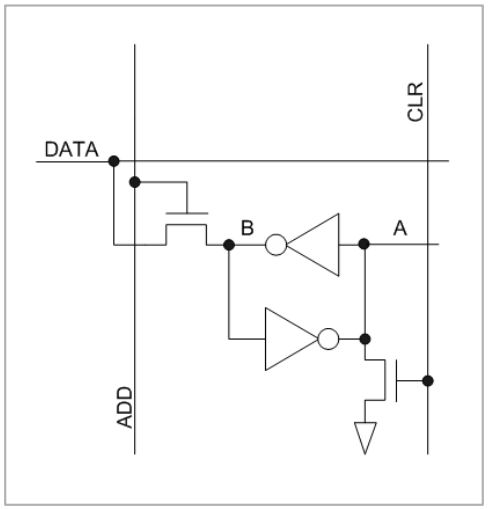
65-nm DP-SRAM SNM Simulation Results

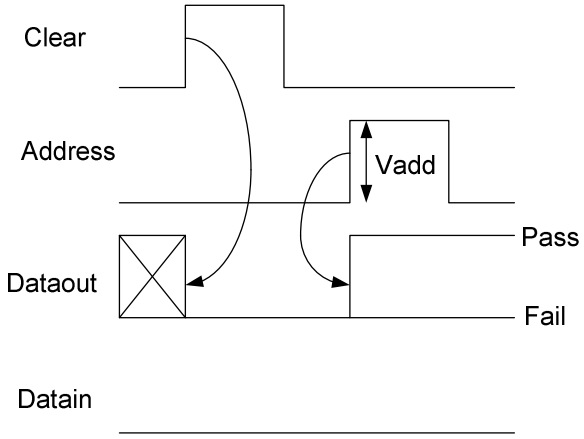


- Accurate modeling of all variation sources required to ensure robust design

Simulation Methodology Validation on 65 nm

Configuration memory cell

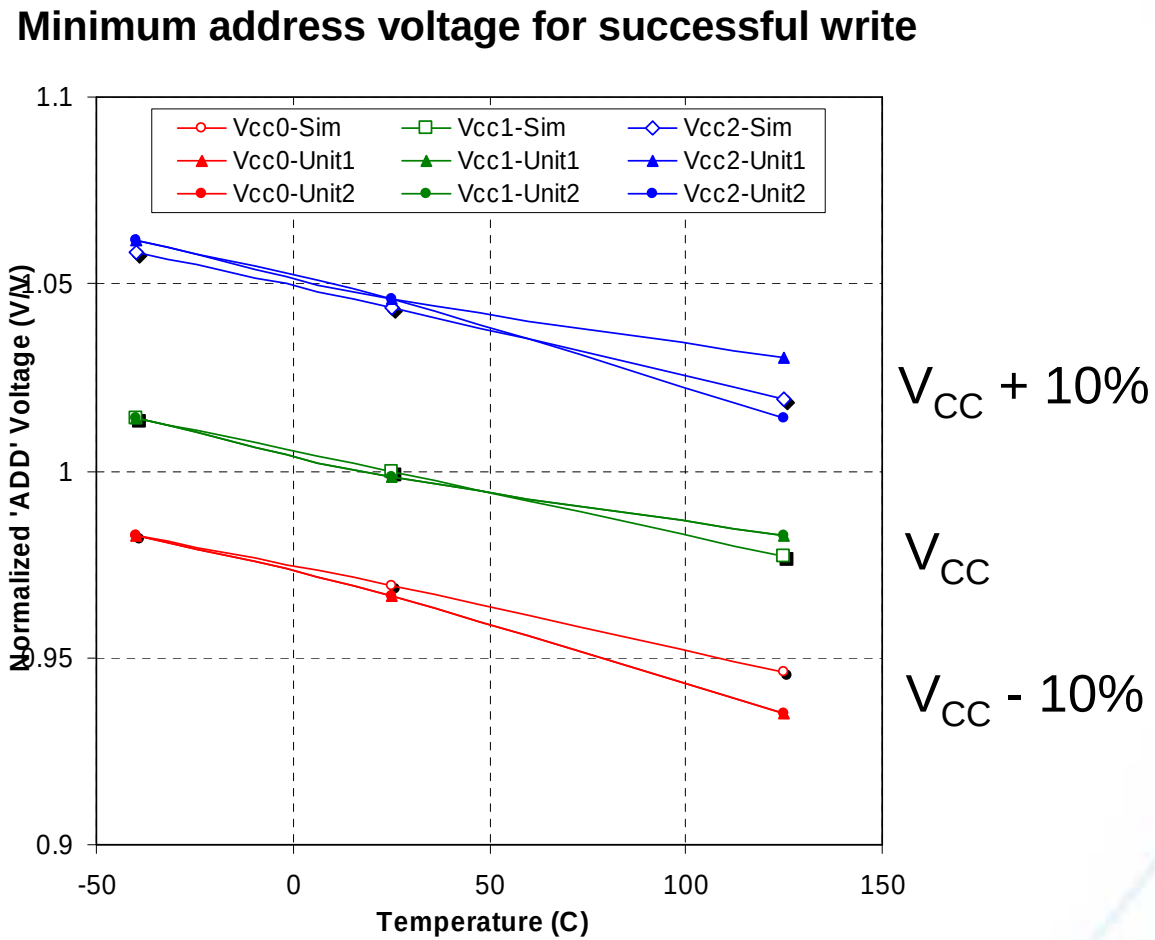




Address voltage	Die yield	
	Simulated	Measured
Level one	1%	1.7%
Level two	99%	98.3%

- Statistical correlation between silicon and simulation for 65-nm FPGA configuration memory
- Simulation used to define 2 address voltage levels
  1. Only 1% die pass, 99% die fail to write
  2. 99% die pass, only 1% fail to write
- 5010 units tested using the simulated address voltage

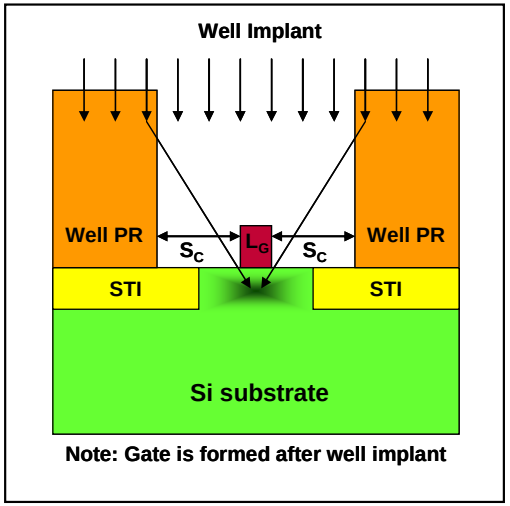
65-nm Simulation vs. Silicon Results



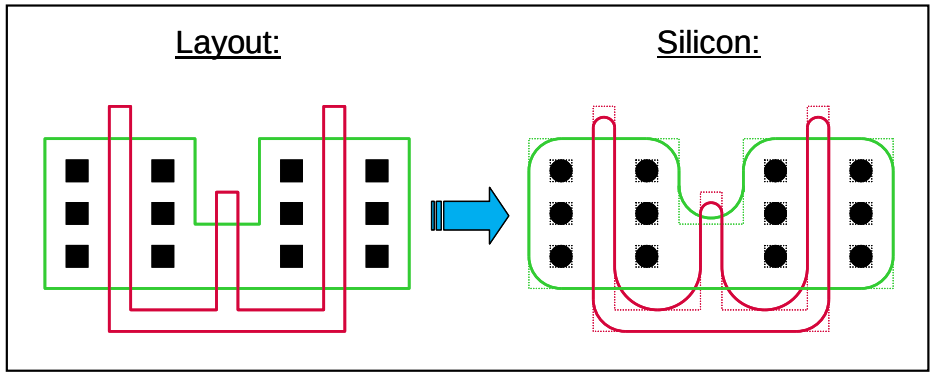
- 2 units at 50% yield distribution compared to simulation
- Temperature and voltage sensitivity validated

Sources of Layout-Induced Variability

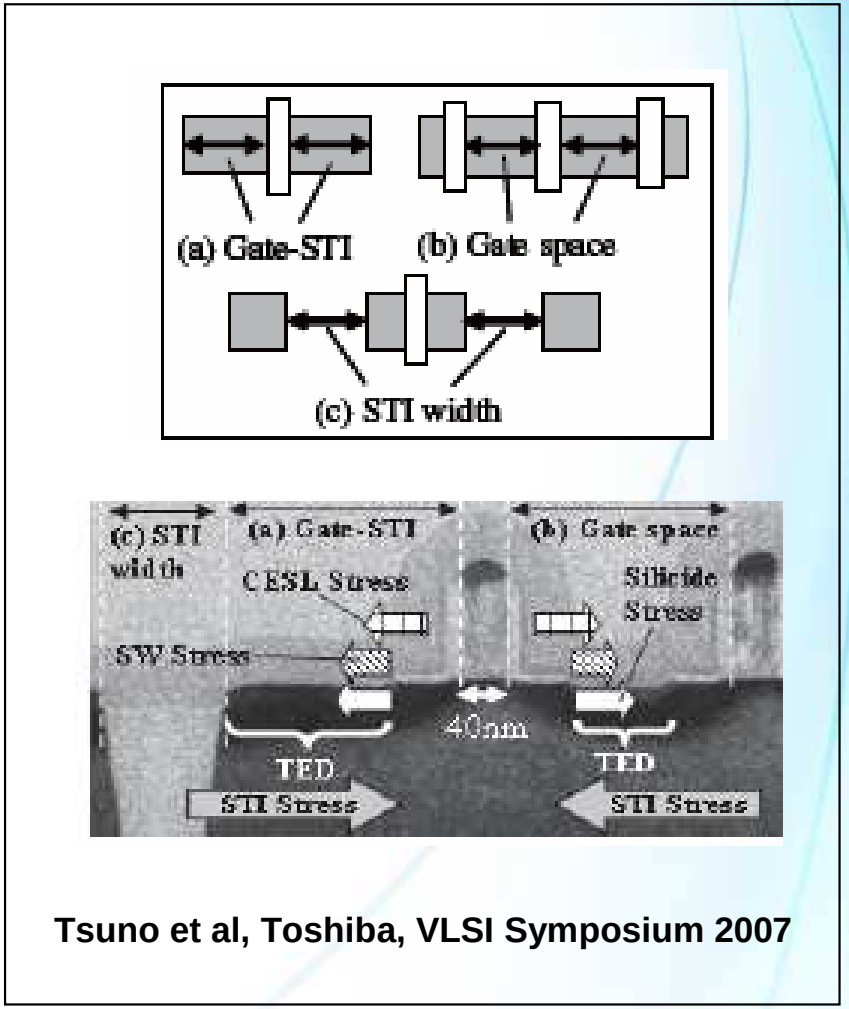
Well proximity effect



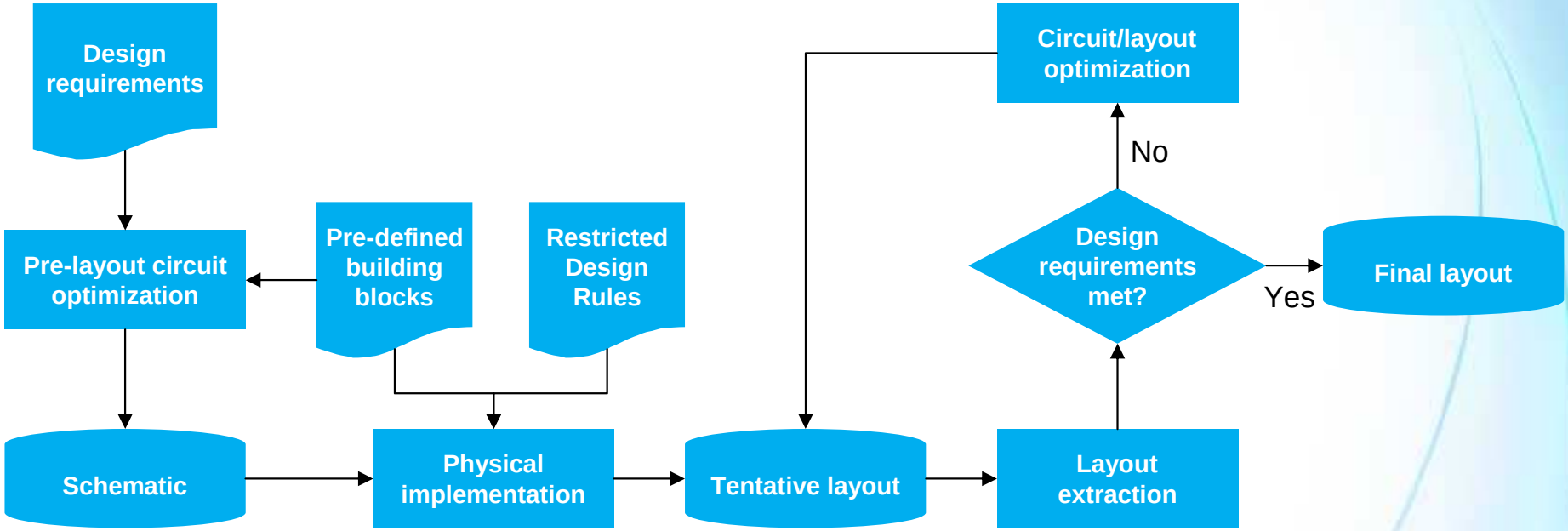
Pattern rounding



Stress effects

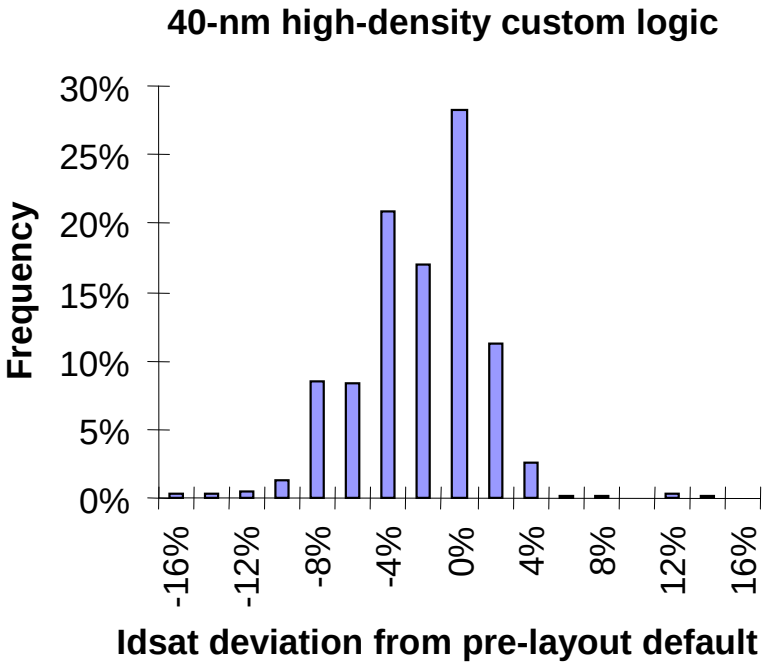


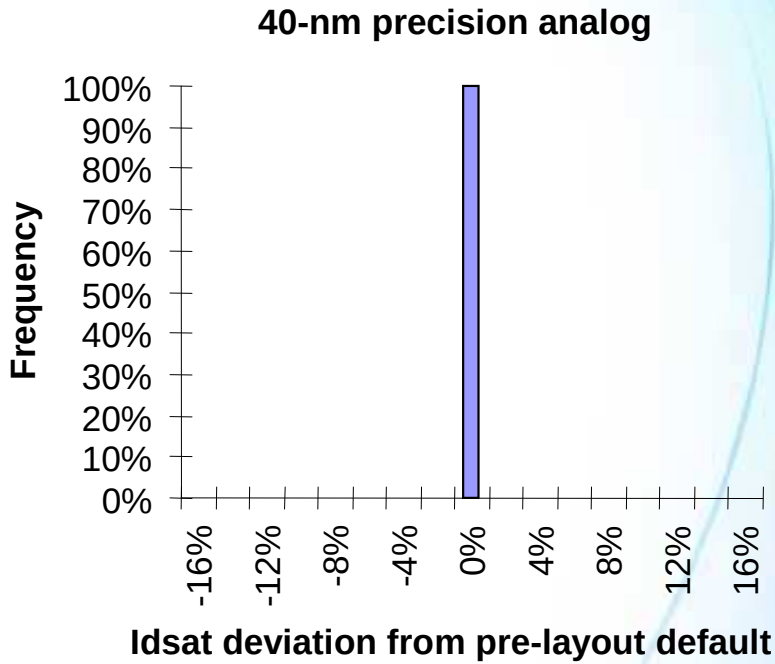
Coping With Layout-Induced Variability



- Restricted design rules defined to limit variability without significant compromise in layout area
- Use of pre-defined building blocks varies with circuit type
  - Logic uses few building blocks to enable high-density custom layout
  - Precision analog uses building blocks extensively to minimize sensitivity to layout effects
- Fast iterations required to optimize for layout-induced variability based on extracted layout parameters

Impact of Layout-Induced Variability

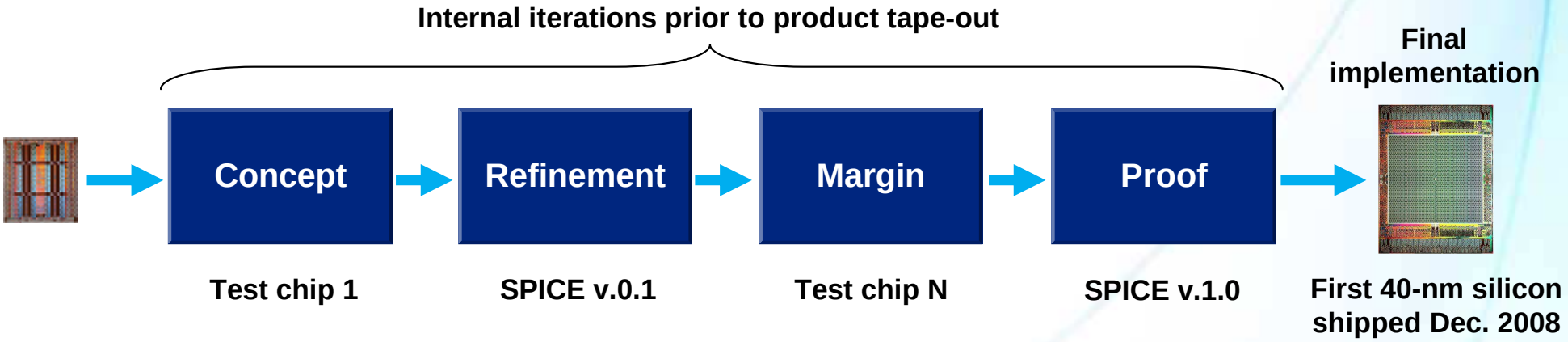




- High-density custom logic is optimized using post-layout simulations to account for systematic layout variability
- Use of pre-defined building blocks for precision analog eliminates deviation between pre-layout and post-layout

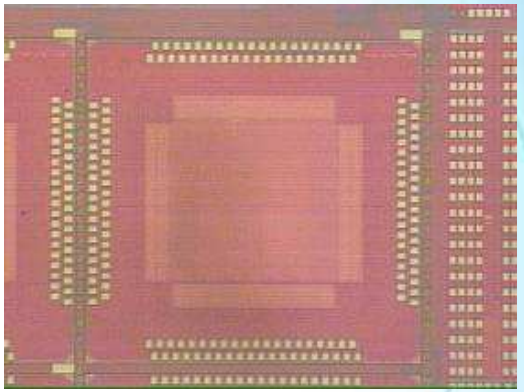
Test-Chip Program

- Test chips combined with accurate models for variability are the foundation of "first silicon to production"
- Test chip goals
  - Develop and verify accurate models for random and systematic variability
  - Reduce risk against process and circuit design uncertainties
  - Validate innovative process improvements and circuit design techniques
  - Optimize tradeoffs for performance and manufacturability
- Test chips progressively increase in complexity as process matures and circuit design nears completion

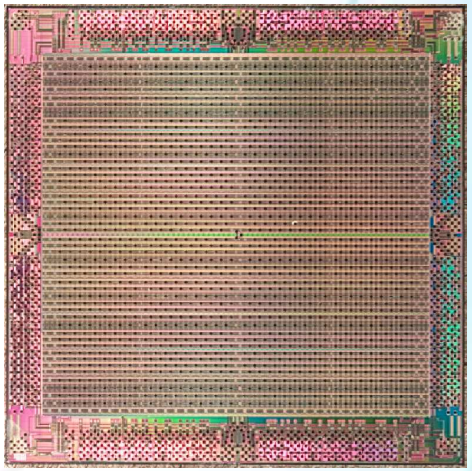


40-nm Development Test Chips

Test chips	Tape out	Major components
TC1	Q4 2005	Early technology assessment, OPC
TC2	Q3 2006	LP process, FPGA fabric, transistors, modeling
TC3	Q4 2006	G process, FPGA fabric, transistors, modeling
TC4-a	Q3 2007	I/Os, logic array block (LAB), analog
TC4-b	Q3 2007	Complete FPGA
TC5	Q3 2007	Transceiver
TC6	Q4 2007	Memory
TC7	Q4 2007	On-chip regulators
TC8	Q1 2008	I/Os, LAB, FPGA fabric



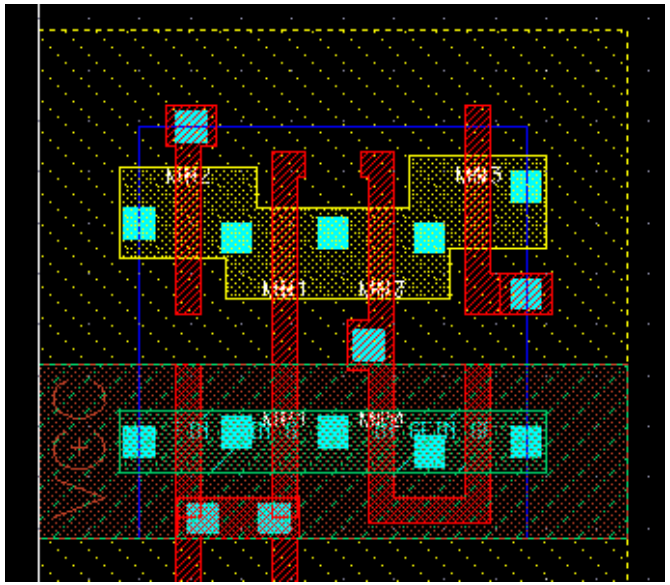
40-nm TC2



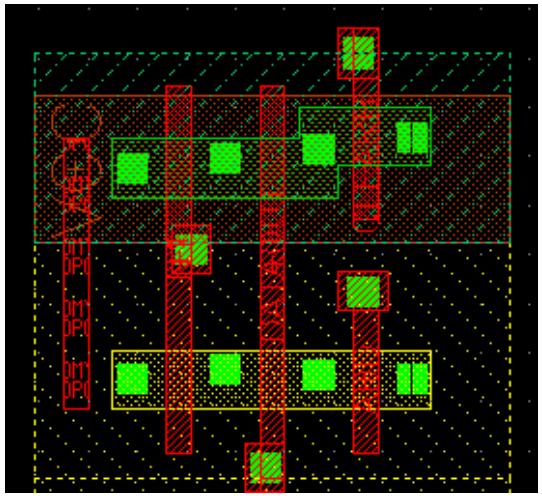
40-nm TC4-b

40-nm Configuration Memory Example

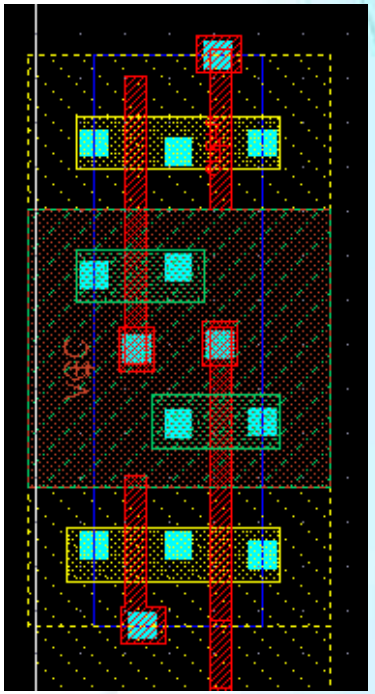
Cell A



Cell B

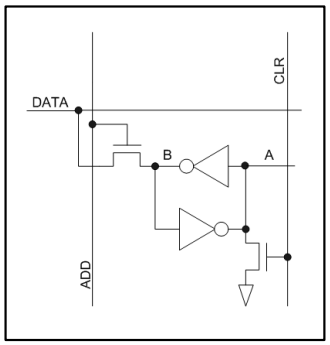


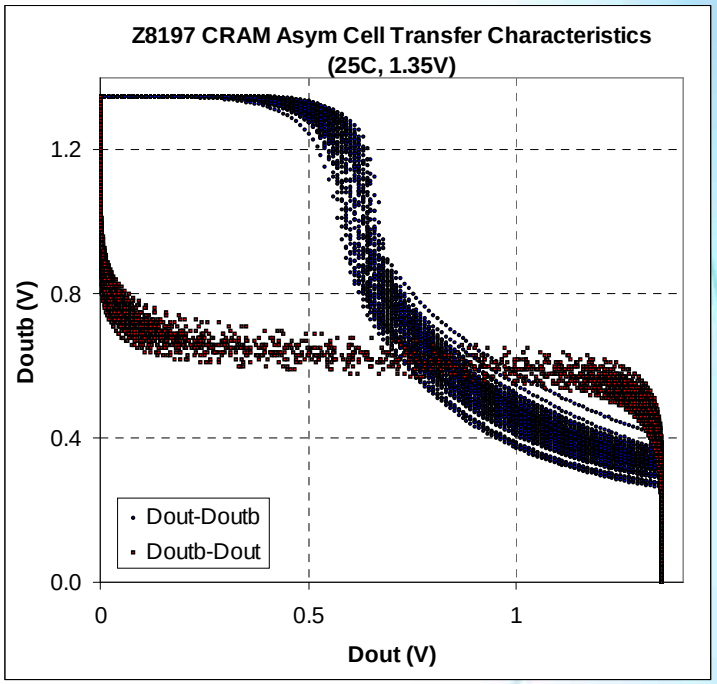
Cell C



- Goals:
  - Develop optimized cell for configuration memory
  - Ensure robust operation for up to 100-Mbit density

40-nm Configuration Memory Results





Cell name	Relative cell area	Static noise margin (v)	
		Median SNM	Std Dev
Cell A	100%	0.210	0.029
Cell B	91%	0.133	0.037
Cell C	82%	0.170	0.036

- Single-ended access for cell read disturb test
- SNM derived from voltage-voltage transfer curve (butterfly chart)
- Cells B and C unstable at 6σ
  - Not suitable for multi-Mb configuration memory without further optimization
- Optimized cell validated prior to product silicon

Variability Outlook

- Variation expected to worsen with scaling
  - Devices are approaching atomic scale
- Device innovations will provide temporary relief
  - Reduced equivalent oxide thickness due to high-k/metal gate
  - Reduced channel doping due to multi-gate transistor structure
- Layout will become more regular to constrain context-dependent variability
- Accurate modeling will be essential
  - Capture all systematic components
  - Robust statistical simulation capability
- Design/architecture innovation will be necessary
  - Variation tolerant circuits (e.g., 8T SRAM)
  - On-chip variation monitoring and compensation
- Close collaboration with foundry will continue to be critical

Summary

- Variability-aware development essential below 90 nm
  - Sub-wavelength lithography, reduced channel dopant count and aggressive process-induced strain significantly increase variation
- Fabless IC development challenges met by collaborative development with foundry partner
- “First silicon to production” approach demonstrated at 65 nm and 40 nm
  - Accurate variation models
  - Robust statistical simulation methodology
  - Test-chip program
- Continuation of collaborative development model and innovative process/design solutions needed for the future