Challenges in Modeling Layout Systematic Effects in Compact Device Models

Andrzej J. Strojwars

PDF Solutions, Inc., San Jose, CA
Carnegie Mellon University, Pittsburgh, PA

MOS-AK/GSA Workshop
Seville, September 17, 2010
Outline

- Motivation
- Variability sources and trends
- Analog/RF Modeling Challenges
- SOC Modeling Challenges
- pdBRIX™ for the 28/20nm technology node
- Conclusions
Technology scaling enablers

- Strain and Substrate engineering remain major performance boosters after HK MG introduction (32 nm IBM Alliance and Intel)
  - Process complexity continues to increase despite (or along with) the benefits from HK-MG

<table>
<thead>
<tr>
<th>Performance booster</th>
<th>Impact on NMOS</th>
<th>Impact on PMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Stress Liner (Gate First)</td>
<td>Mobility improvement from Tensile liner</td>
<td>Mobility improvement from Compressive liner</td>
</tr>
<tr>
<td>eSiGe</td>
<td>N/A</td>
<td>Mobility improvement from Compressive stress</td>
</tr>
<tr>
<td>Contact-induced strain</td>
<td>Bar contacts with recessed NMOS Active</td>
<td>N/A</td>
</tr>
<tr>
<td>Gate induced strain (Metal Gate Last)</td>
<td>N/A</td>
<td>Additional Strain allowed by Replacement Gate</td>
</tr>
<tr>
<td>SMT - Stress Memorization (Gate First)</td>
<td>Mobility improvement due to Strained Poly – induced stress</td>
<td>N/A</td>
</tr>
<tr>
<td>S/D emb SiC</td>
<td>Mobility improvement from longitudinal Tensile stress</td>
<td>N/A</td>
</tr>
<tr>
<td>Substrate orientation</td>
<td>(100) Beneficial across wide range of stress level</td>
<td>(110) better for hole mobility, gain deteriorate with stress level</td>
</tr>
</tbody>
</table>
Outline

- Motivation
- Variability sources and trends
- Analog/RF Modeling Challenges
- SOC Modeling Challenges
- pdBRIX™ for the 28/20nm technology node
- Conclusions
Variation Trends: Total Variation

- Identical structure data presented here (no layout variability impact included)
- Total variation is increasing with scaling
- At 32 nm node, narrow transistors have $3\sigma$ variability > 45%, diminishing the benefits of scaling
Random Variability

- Random variations play an important role
  - Line edge roughness (LER)
  - Random dopant fluctuations (RDF)

- Metal gate/High-K saves us for one generation only: 32nm/28nm

- RDF impact can be minimized by choosing different transistor architectures (FinFET, Ultra Thin Body or Fully Depleted SOI)
32/28nm logic has 2000+ unique transistors to characterize & model
45,000+ DOE levels required to characterize transistor models
Assuming 8 replicas and Kelvin/leakage arrays more than 400k transistors needed in a test chip for full characterization
### Layout Effects in Nanometer Technologies

- **Printability and stress effects increase systematic layout variation**

<table>
<thead>
<tr>
<th>Systematic variation (**)</th>
<th>Root causes</th>
<th>Impact at the 45/32nm technology node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate poly orientation</td>
<td>Poly lithography</td>
<td>N/A</td>
</tr>
<tr>
<td>Gate poly pitches</td>
<td>Poly OPC and lithography, Stress Layers, Loading effects from Etch and film deposition</td>
<td>3-5% - printability, 15% - poly pitch related stress, 5% - contact space related stress</td>
</tr>
<tr>
<td>Poly corner rounding</td>
<td>Poly OPC and lithography</td>
<td>15%</td>
</tr>
<tr>
<td>Transistor location in multi-gate transistor</td>
<td>Poly patterning and stress impacted by local neighborhood differences</td>
<td>15%</td>
</tr>
<tr>
<td>Active corner rounding</td>
<td>OPC and lithography, stress</td>
<td>7%</td>
</tr>
<tr>
<td>Un-modeled narrow width effects</td>
<td>Stress effects, poly step height</td>
<td></td>
</tr>
<tr>
<td>Gate to active edge</td>
<td>STI stress, e-SiGe stress</td>
<td>10-20% for PMOS with e-SiGe</td>
</tr>
<tr>
<td>Well proximity</td>
<td>Implant scattering</td>
<td></td>
</tr>
<tr>
<td>Nwell-Pwell separation</td>
<td>Gate counterdoping and misalignment</td>
<td></td>
</tr>
<tr>
<td>Contact density and placement</td>
<td>Silicide sheet resistance and stress layers</td>
<td></td>
</tr>
</tbody>
</table>
Motivation

Variability sources and trends

Analog/RF Modeling Challenges

SOC Modeling Challenges

pdBRIX™ for the 28/20nm technology node

Conclusions
# Transistor Care-abouts vs Applications

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Digital Logic</th>
<th>Digital SRAM</th>
<th>Baseband Analog</th>
<th>RF (&lt; 5GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDSAT</td>
<td></td>
<td>Critical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOFF</td>
<td>Critical</td>
<td></td>
<td>Somewhat critical (SNR degradation)</td>
<td>Not critical</td>
</tr>
<tr>
<td>VTSAT</td>
<td></td>
<td>Critical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Igate</td>
<td>Critical</td>
<td></td>
<td></td>
<td>Impacts I&lt;sub&gt;d&lt;/sub&gt;s matching, SNR</td>
</tr>
<tr>
<td>C&lt;sub&gt;gg&lt;/sub&gt; (C&lt;sub&gt;ovl&lt;/sub&gt;+C&lt;sub&gt;gb&lt;/sub&gt;)</td>
<td>Critical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C&lt;sub&gt;j&lt;/sub&gt;</td>
<td>Critical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm at low I&lt;sub&gt;d&lt;/sub&gt;s</td>
<td>Not critical</td>
<td></td>
<td>Critical</td>
<td></td>
</tr>
<tr>
<td>G&lt;sub&gt;d&lt;/sub&gt;s</td>
<td>Not critical</td>
<td></td>
<td>Critical</td>
<td></td>
</tr>
<tr>
<td>Mismatch</td>
<td></td>
<td>Show stopper</td>
<td></td>
<td>Critical</td>
</tr>
<tr>
<td>Linearity (IP3)</td>
<td>Not Critical</td>
<td>Somewhat critical</td>
<td>Critical</td>
<td></td>
</tr>
<tr>
<td>f&lt;sub&gt;T&lt;/sub&gt;, f&lt;sub&gt;max&lt;/sub&gt;</td>
<td>Not critical</td>
<td></td>
<td></td>
<td>Critical</td>
</tr>
<tr>
<td>Noise (1/f, NF&lt;sub&gt;min&lt;/sub&gt;)</td>
<td>Not critical</td>
<td>Critical (SNR degradation)</td>
<td>Critical</td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;sub&lt;/sub&gt;</td>
<td>Low R&lt;sub&gt;sub&lt;/sub&gt; for latchup</td>
<td></td>
<td></td>
<td>High for Q</td>
</tr>
</tbody>
</table>

- Analog/RF designs have different/additional care-about from digital logic and SRAM technology drivers
Consequence of Transistor Scaling for Analog/RF

<table>
<thead>
<tr>
<th>Care-about</th>
<th>45/40nm</th>
<th>32 nm MG/HiK</th>
<th>32nm Poly/SiON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic gain (gm/gds) (Lmin, Vgs=Vds=Vdd/2)</td>
<td>~10</td>
<td>Similar or better due to gm ↑ and (potentially) gds ↓</td>
<td>Similar to or better than 40nm due to gm</td>
</tr>
<tr>
<td>Gate Leakage</td>
<td>~0.75 nA/um²</td>
<td>25-1000X lower</td>
<td>Same or higher than 45 nm</td>
</tr>
<tr>
<td>Mismatch</td>
<td>nMOS Avt ~ 3.0 mV/um</td>
<td>Significantly better than 40 nm &amp; Poly/SiON</td>
<td>Avt same as 40 nm, min area matching worse</td>
</tr>
<tr>
<td>Linearity</td>
<td></td>
<td>Degraded, but potentially less compared to Poly/SiON</td>
<td>Degraded due to scaling</td>
</tr>
<tr>
<td>1/f noise</td>
<td>~1e-11 V².um²/Hz</td>
<td>Same as 45/40</td>
<td>Potentially degraded due to nitridation</td>
</tr>
<tr>
<td>f_T/f_max</td>
<td>~370/410 GHz</td>
<td>Higher than 45/40</td>
<td>Higher than 45/40</td>
</tr>
<tr>
<td>Layout dependence</td>
<td>Severe</td>
<td>Similar to Poly/SiON</td>
<td>Similar to 40nm HP</td>
</tr>
<tr>
<td>NF</td>
<td></td>
<td>Similar to Poly/SiON</td>
<td>Improved by scaling</td>
</tr>
<tr>
<td>Thermal Stability</td>
<td></td>
<td>Potentiality worse</td>
<td>Same as 40nm</td>
</tr>
<tr>
<td>Reliability (NBTI/HCl)</td>
<td>&gt; 10 years</td>
<td>Same as Poly/SiON (&gt;10 yrs)</td>
<td>Potentially degraded due to nitridation</td>
</tr>
<tr>
<td>Process cost/other risks</td>
<td>Std CMOS cost</td>
<td>Double patterning</td>
<td>Double patterning</td>
</tr>
</tbody>
</table>
Transistor Scaling Impact Summary

- The 32/28 nm technology node has distinct technology choices, each with its advantages and risks
  - In addition, they share some common risks, like increased layout sensitivity

- All the consequences of these choices on Analog/RF design are unclear
  - The trade-off is more clear for digital logic design and SRAM

- Poly/SiON
  - Next generation of a proven technology; better understood trade-off in terms of performance and cost
  - In general, devices will have worse electrostatics compared to MG/HiK, with its attendant impact on Analog/RF design (linearity, matching, leakage)
  - Reliability degradation as a consequence of Tox scaling

- Metal Gate/Hi-K
  - Benefits of devices with much better electrostatics (matching, linearity, leakage)
  - Improved reliability due to larger Tox(physical) ⇒ Lower E-field
  - However, new phenomena like Thermal stability, Resistor TC
  - Process cost

- Accurate and comprehensive characterization will be essential for effective utilization of either of the technology choices for Analog/RF design
Analog/RF Design in 40nm and 32nm Technology Nodes

- **40nm Process** – the best analog/RF process since 130 nm [Klaas Bult - Broadcom Analog/RF CTO at 2009 ESSDERC Panel Discussion]
- All analog devices use minimum L (same as digital)
- Large area (equivalent width) devices used to average RDF effects
- Full dummification: very few layout patterns
- Mismatches modeled precisely based on Si characterization results
  - Small number of patterns to characterize
- Very aggressive shrink achieved
- No area penalty due to dummification

- **32/28nm processes with MGHK:**
  - Improved transistor matching

- Intel uses only 2 L values in their SOC’s: one for digital and one of analog transistors [Mark Bohr - Intel at 2009 ESSDERC Panel Discussion]

- Full dummification to eliminate layout systematics
Outline

- Motivation
- Variability sources and trends
- Analog/RF Modeling Challenges
- SOC Modeling Challenges
- pdBRIX™ for the 28/20nm technology node
- Conclusions
Layout Dependency Increasingly Important

45/40nm PMOS Measurement Layout Dependency

- Performance dependence on layout increasing
- Model accuracy decreasing

Key is to evaluate wide range of layout dependencies relevant to product
Process advancements have made device behavior highly sensitive to the drawn geometry.

Modern day modeling methodology fails to accurately model entire design space.

Restricting the design space with regular fabrics can provide more accurate models.
Device Modeling of Layout Interactions

![Graph showing normalized IdsAT/μm for nFET and pFET with varying gate width. The graph indicates the effects of different extensions (Short Ext, Inter. Ext, Long Ext) on the normalized IdsAT/μm for both nFET and pFET.]
Device Modeling of Layout Interactions
Variability: Is this a problem designers should be addressing?

- **Designers’ historical perspective:**
  - “My world starts with SPICE models and Design Rules. It is up to the Foundry to get the SPICE models correct and the support the entire valid DR space”

- **Current reality: silicon**
  - Even typical gridded SOC designs can have >10,000 different transistor patterns
  - Complexity of interaction effects makes it hopeless for foundries to cure this problem
  - Physics won’t change or improve as node matures – certain patterns will always have higher variability

- **Let’s restore designers freedom at advanced nodes**
  - Selecting right set of low variability patterns
  - Validating SPICE models in silicon for those patterns
  - Ensuring manufacturability of metal patterns used in design
Need for Limiting Number of Patterns

- Improved predictability requires limiting the number of unique patterns
- New patterns are created from cell abutments
- Without wavelength scaling the optical interaction remains constant
  - More cells fall within the same optical region of influence

Problem pattern

Pattern influence range for 193nm lithography
Motivation

Variability sources and trends

Analog/RF Modeling Challenges

SOC Modeling Challenges

PDRIX™ for the 28/20nm technology node

Conclusions
pdBRIX™ Logic Templates

- Templates: Limited number of single-stage logic topologies (including all primitives: INV, NAND, etc.)
  - We select approximately 70 of them that are most useful across wide application range given transistor stack height constraints

- Choosing limited set of layout constructs to implement templates provides pattern predictability
  - Well chosen layout constructs can facilitate exhaustive qualification of all possible patterns including neighborhoods

- All logic cells and functions can be constructed as combinations of the qualified template layout set

Reduce variability through pdBRIX™ design
Exhaustively characterize all patterns used
pdBRIX™ Library Creation Infrastructure

**STANDARD CELL LIBRARIES:**
Complete standard cell library assembled from templates

Silicon predictability achieved by:
- Merging, embedding, and abutting silicon qualified templates
- pdBRIX ensures zero new silicon layout patterns are created
- Tighter (simplified) SPICE models

**TEMPLATES:**
Small set of single stage logic functions built on Fabrics

Co-optimized with fabrics and library:
- Functions selected for optimal library coverage
- Layouts co-optimized with fabrics for best area, power, and performance
- Minimize number of patterns created

**FABRICS:**
2-D grid environment defining allowed poly and metal shapes

Selected based on:
- Litho printability
- SRAM compatibility
- Yield optimization
- Product needs such as redundancy
Controlling layout patterns enables exhaustive qualification of the layout space over the relevant interaction range.

Fewer Patterns $\rightarrow$ Better Predictability

- Construction by templates
- Pattern regularity
- 30 gridded cell lib
- Non-gridded cell lib trajectory

Controlling layout patterns enables exhaustive qualification of the layout space over the relevant interaction range.
32/28nm Bulk Fabric Fabricated

Fabric chosen to accommodate manufacturing objectives and template requirements for design

Fabric characteristics

- Fully gridded transistors
- Limited diffusion jogs
- Relaxed metal pitches
- Limited “two-way” metal patterns
- 100% active contact redundancy (9T+)
- Portable to 28nm

Qualifying layout constructs

- 8T, 10T and 13T silicon qualified
- 9T, 11T, 12T and 14T+ available via virtual qualification

X = !( E(A+B)(C+D) + F )
Fabric Optimization: Parametric Variability Control

- Limited set of regular design can eliminate over half of NMOS Idsat variability due to poly effects
- More predictive transistor performance in std. cell & IP layout
Transistor layout and layout neighborhood can significantly impact transistor performance

~35% reduction in Idsat variability due to layout & neighborhood for pdBRIX NMOS transistors

pdBRIX limits total number of transistor logic patterns
- Can avoid high variability patterns
- Enables more accurate Si-to-SPICE matching of limited layout patterns
~30% reduction in \( I_{dsat} \) variability due to layout & neighborhood for pdBRIX PMOS transistors

- pdBRIX transistors can be qualified & modeled over all patterns. Cannot insure performance over all DR-compliant patterns in an SoC
... Enabling Design With Low Variability

CDF of $|\Delta I_{dr}|$ vs. Reference

<table>
<thead>
<tr>
<th>$%$</th>
<th>CDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>1</td>
</tr>
<tr>
<td>75%</td>
<td>0.75</td>
</tr>
<tr>
<td>50%</td>
<td>0.5</td>
</tr>
<tr>
<td>25%</td>
<td>0.25</td>
</tr>
<tr>
<td>5%</td>
<td>0</td>
</tr>
</tbody>
</table>

Source PDF Solutions

Significantly tighter transistor performance distribution possible from limiting transistor neighborhoods

- Templates-based
- DR-based
Achieves 22/20nm Pattern Control

Pattern Count vs. Block Area
22/20nm node, 2 pitch interaction range

- Templates provide pattern constraint beneficial for source-mask optimization algorithms

Source PDF Solutions
Outline

- Motivation
- Variability sources and trends
- Analog/RF Modeling Challenges
- SOC Modeling Challenges
- pdBRIX™ for the 28/20nm technology node
- Conclusions
pdBRIX Templates Improve Model - Hardware Correlation

- **Templates: “Model = Hardware” → helps achieve working first silicon**
  - New simplified interface between foundry and std. cell / IP designers
  - Templates plus basic (process-integration driven) design rules replace increasing complex design rule manuals
  - Results in “correct by construction” IP
  - Enable compatibility across foundries

- **Templates achieve “Model = Hardware” from:**
  - Comprehensive Si verification enabling foundry to focus on specific patterns
  - Low layout pattern count compatible with SMO / DPT
  - Elimination of product-specific “hot spots” or yield loss mechanisms
  - Limited transistor patterns enable complete Si-to-SPICE validation

- **Templates have been validated:**
  - 40nm: adopted by Toshiba
  - 32/28nm: validated on IBM Alliance CV’s
  - 22/20nm: validated with ASML/Brion Tachyon SMO (Si verification underway)
  - Fabless: adopted at 20nm
Conclusions

- Variability crisis continues although MGHK architecture helps for one generation (32/28nm)
- Si-SPICE mismatch has become a key limiter
- Systematic layout effects must be eliminated
- Proposed comprehensive methodology for design-process co-optimization to reduce the cost per good die
  - Extensively verified in silicon
- Regular design methodology with limited number of patterns only viable means for advanced lithography solutions such as SMO and interference assisted lithography
- Current methodology can be used to concurrently determine the optimal lithography and design solutions for application domain
  - Successfully Implemented in the 22/20nm technology node