DFM reality in sub-nanometer IC design

Nishath Verghese, Philippe Hurat
Clear Shape Technologies

Abstract:
The impact of sub-nanometer (below 90nm) effects on IC designs needs to be clearly understood to ensure that (1) manufacturing variations are considered during design to avoid catastrophic failures, and (2) the expected performance simulated in design is actually realized on silicon to avoid parametric failures.

This paper discusses design for manufacturing solutions that enable designers to predict systematic manufacturing variations during design to detect and repair catastrophic and parametric failures. This paper presents real examples of design sensitivities to sub-nanometer manufacturing variations and the need to correctly analyze, optimize and verify the design before manufacturing by using appropriate EDA solutions which bring the effects of manufacturing variations in the design flow.

I. Introduction

Historically, global random variations have dominated local systematic variations. With the introduction of sub-90 nm semiconductor technologies, primarily centered around resolution problems, copper/low-k interconnects and CMP, local systematic variations have become decisively greater than global random variations. In practice, this means that what is modeled and designed based on ideal layout is not what's seen on silicon. The result is large margins applied at every stage of the design, long timing closure cycle time, long manufacturing turnaround times, catastrophic and parametric yield loss and major bottlenecks at the RET/OPC stage.

To address these sub-nanometer effects, manufacturing teams are using resolution enhancement techniques (RET) to compensate for the use of 193nm steppers and to predict the shape accuracy on wafer. The more well known of these techniques used in production today, include strong or weak Phase Shift Mask (alt-PSM and att-PSM), Optical Proximity Correction (OPC), Assist Features (AF) and retargeting.

The problem is that the design engineer has no control of the RET results and its impact on the expected results. The results are the all too familiar delays (multiple design, manufacturing passes) in product silicon qualification. These delays are due to design modifications or engineering change orders (ECO), which can be avoided and almost eliminated if a design engineer had the ability to determine design sensitivity to sub-nanometer manufacturing effects.

Even if proper design for manufacturing methodology is used to avoid systematic catastrophic failures, manufacturing variations can still severely impact the performance of the design, causing parametric failures.

In this paper we present the real examples of design sensitivities to sub-nanometer manufacturing variations and elaborate on the need to correctly analyze, optimize and verify the design before manufacturing by predicting and correcting the effect of manufacturing variations on the design yield and performance.

II. Sub-nanometer Variations

Sub-nanometer variations can be classified into two categories: *random* variations and *systematic* variations. Random variations or effects are inherent fluctuations in process parameters, such as random dopant fluctuations, that have no known deterministic explanation and cannot be influenced by design.

Systematic variations depend on the layout shapes and are deterministic and predictable when modeled correctly. The main sources of systematic variations include optical proximity effects and OPC/RET residual error, stress and CMP dishing. Each has a predictable impact on the performance of a transistor gate, and the thickness and shape of an interconnect wire.

The primary sources of manufacturing variation (ΔR, ΔC, ΔCD) include:

- Deposition
- Etching
- Sub-wavelength lithography
- CMP

As dimensions shrink, physical effects that used to be insignificant become a more significant fraction of the whole and can no longer be considered as noise.

1-4244-0630-7/07/$20.00 ©2007 IEEE.

The complexity of sub-nanometer semiconductor electrical elements such as CMOS transistors creates unexpected and unwanted timing, noise and power results. The impact is poor semiconductor product yields and reliability. In order to correct and compensate for these anticipated poor results, a classification of sub-nanometer effects is required. The classification is broken down into areas so that design tools and methods can be applied to achieve the maximum improvement in each area. Each sub-nanometer effect classification is described below.

A. Variability in complex shapes on silicon

Figure 1 shows the increasing complexity of sub-nanometer shapes in silicon. Five or more interconnect layers using multiple materials, tens of millions of gates and highly sensitive analog circuits are included in today's and future sub-nanometer SOCs.

Figure 1: Sub-nanometer image of copper interconnect

Not only is the final shape not the same as the drawn shape, but there is "variability" in the length and width of both transistors and interconnect. The DFM challenge is to know what shapes are sensitive to manufacturing variations, how they impact timing, power and noise yield. A comprehensive understanding of all the undesired shapes is very complicated to achieve as non-linear 2D effects make it impossible to capture all undesired patterns with a closed set of design rules.

B. Sub-90 nanometer device variation

Transistor performance depends heavily on gate dimension. Small gate length variations can result in significant variation in on-state current, I_{on} and off-state current, I_{off}. With shrinking gate dimensions, dependence of transistor current is increasingly non-linear in channel length. As a result, the variability in I_{on} and I_{off} has been increasing with process node size, as shown in Figure 2 which illustrates the variation in I_{on} and I_{off} for a 90nm high performance device with a final, post etch channel length of 50nm. It is evident that a 10% change in channel length can cause a 40% variation in I_{on} and 90% variation in I_{off}.

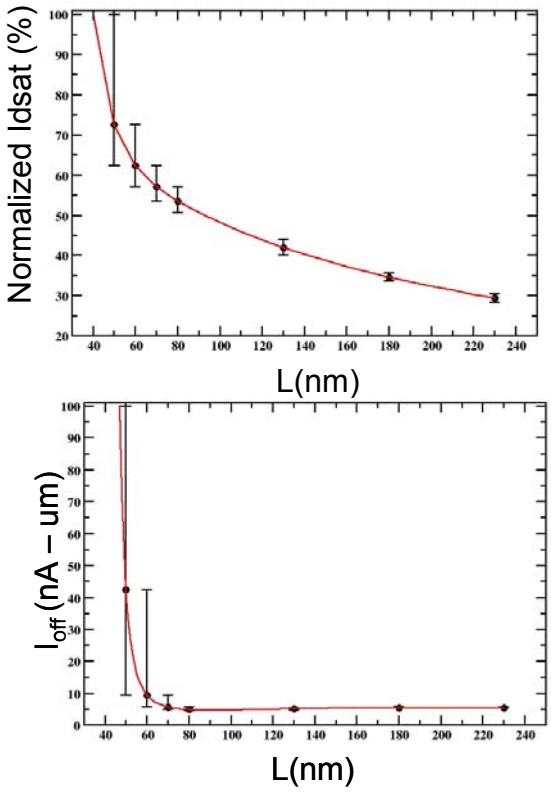

Figure 2: I_{on} and I_{off} variation due to transistor length

Since delay is directly proportional to I_{on}, a 10% transistor channel length variation can translate to -15% to +25% gate delay variation, as shown in Figure 2.a. The impact of variability on leakage power is well known [1] and a mere 6% variability in channel length has been known to cause product failures [2]. This problem is getting worse at 65nm and below where systematic variations of 3nm on a transistor gate can cause a 20% variation in delay and have a 2x impact on leakage power.

Because of the sub-nanometer variations in transistors, it is impossible to accurately predict final silicon performance without modeling final silicon shapes and their associated electrical behavior. Without accurate prediction of the systematic variations inherent to nanometer-scale, it is impossible to be confident in manufacturing consistency and overly pessimistic guard bands have to be used during design.

To cope with the increasing sensitivity of transistors to manufacturing variations, a device model must account for the electrical variation associated with manufacturing variations. For consistent silicon success

and more aggressive designs, variations must be predicted and taken into account early into the design process.

In the analog world where the expected performance is for two transistors to be matching in electrical behavior, these sub-nanometer effects amplify the condition known as mismatch.

C. Transistor mismatch

Transistor mismatch is the condition when two transistors have the expectation of being electrically equivalent. In analog and RF applications, this means that when the two transistors or interconnects are not the same, overall performance does not meet target specifications. To correct the problem, large silicon area is taken up by circuitry designed to compensate for geometric width and process variations.

Figure 3 shows the geometric dependency of transistor V_t mismatch where WG and LG represent the width and length of the transistor respectively and VB is the body bias [4]. The four geometric quadrants range is from large width, large length to minimum width, minimum length. The minimum width transistors do not match, creating a broad scattering in threshold voltage (V_t) measurements. This means that small transistors have a high likelihood of being mismatched or different. This sub-nanometer effect is related to the variations in I_{on} and I_{off} shown in Figure 2 through the DIBL (Drain Induced Barrier Lowering) effect [3].

A significant reason for the variation in width and length between matched transistors is the systematic defects due to lithography and etch. To compensate for this systematic defect, the transistor width is increased at the penalty of using a larger amount of area which increases die size.

Circuit designers account for mismatch by performing statistical simulations that predict the performance variation due to manufacturing variations. However, in creating these statistical models, care must be taken to abstract out the systematic components from the device parameters so that only the truly random components are modeled statistically. The transistor's mismatch sensitivities are identified and reported to the design engineer so that design improvements can be made to ensure that the transistors match better. Common compensation techniques are to increase the matched pair size and change the transistor architecture to a common centroid layout as shown in Figure 4.

Source: Freescale (P. Drennan) WESCON 2005

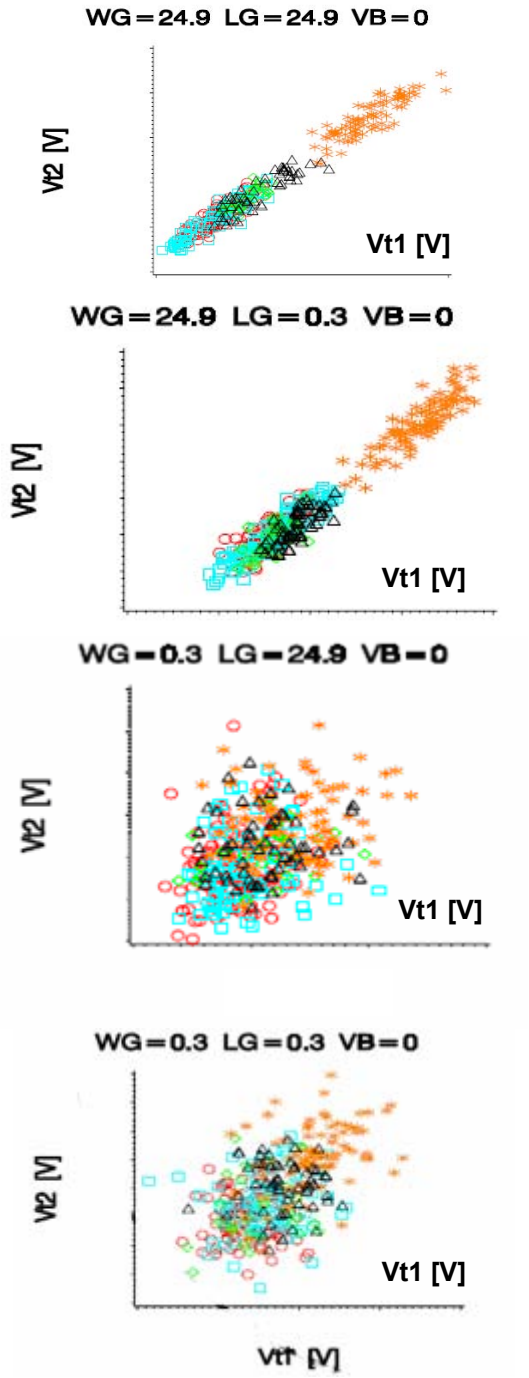

Figure 3: Transistor matching compared to gate length and width [4]

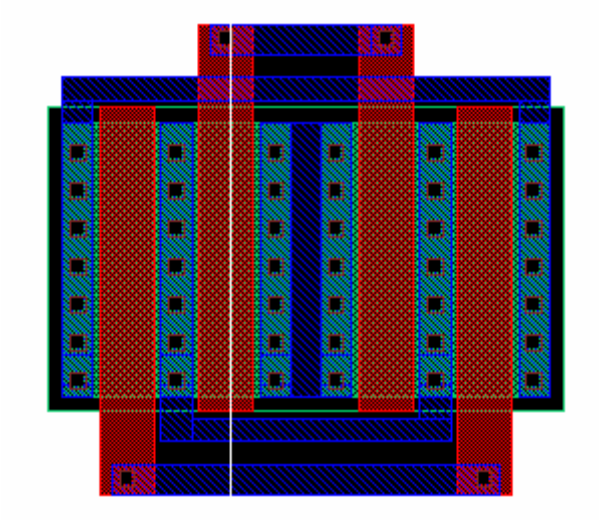

Figure 4: Common centroid transistor layout configuration

D. Sub-90 nanometer interconnect variation

Interconnect parasitics are significant and complex components of circuit performance, signal integrity and reliability for sub-nanometer circuits. Copper/low-k process variability is becoming increasingly important to accurately model and understand. Sub-90nm interconnects with their narrow and tall (Z-plane) configuration create variations in the Z-plane, which, when added to the X-Y variations due to lithography and etch, create significant interconnect RC variability.

Linewidth variations are amplified because copper and aluminum interconnect sheet resistance varies as a nonlinear function of line width, as shown in Figure 5). Copper resistivity increases dramatically below 90 nm due to increased electron scattering on grain boundaries and interfaces [5]. This phenomenon causes a significant systematic change in sheet resistance as a function of line width. If line width cannot be predicted accurately, the result is inaccurate electrical and timing analysis.

Figure 6 shows the nonlinearity of RC delay with respect to interconnect shapes. The graph shows that between normalized values of 0.2 and 0.4, interconnect width has a dramatic impact on interconnect delay. Comprehending a circuit's sensitivity to interconnect RC variation and applying a model for the same in electrical and timing analysis enables the design engineer to make the necessary design adjustments to control timing variability of the design. This understanding can be leveraged to make more manufacturable and stable designs by adjusting length, width and spacing of shapes.

Source: N.S.Nagaraj, ISQED 2003

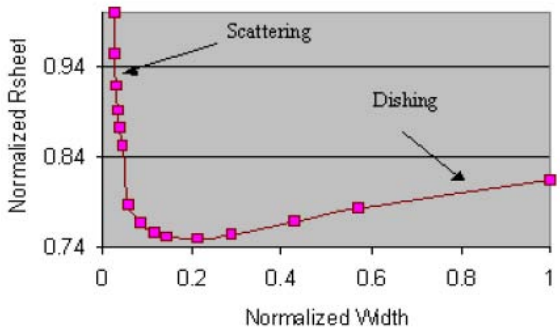

Figure 5: Interconnect resistance variation due to width [5]

Source: TI (Bittlestone) – ISSCC 05

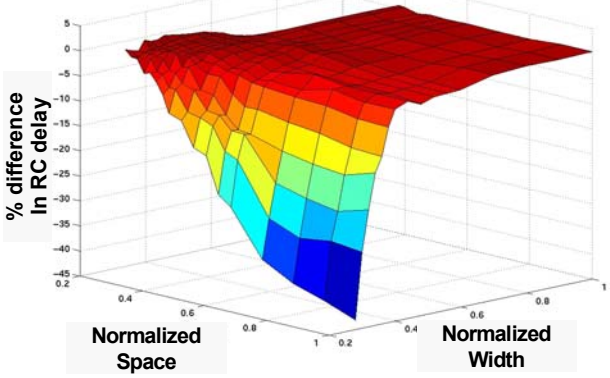

Figure 6: Interconnect RC delay variation due to normalized width and spacing [5]

III. Current Models for Sub-nanometer Variability

The electrical and timing models representing interconnect RC, transistor I_{on} and transistor I_{off} variation are essential to comprehending a circuit's sensitivity to sub-nanometer effects. Guard-bands are added to the SPICE corner models and rule-based derating of gate delays is employed in static timing analysis in an attempt to circumvent transistor variability. The SPICE corner models are derived from idealized test structures on silicon that do not reflect possible systematic shape variations due to layout context. As shown in Figure 7, the +/- 3 sigma corners derived from measurements sampled on multiple dies and wafers are further guard-banded to attempt to capture systematic variations. This however, results in overly conservative corner models. In the absence of better predictability of the impact of systematic shape variations, such guard bands and timing margins are the

only resort in today's flows. Furthermore, if these guard-bands are not large enough, timing failures will go undetected.

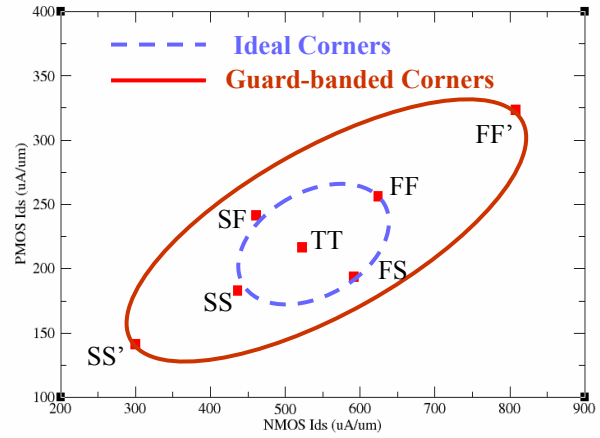

Figure 7: SPICE corner models guard-banded to account for systematic shape variations.

Interconnect modeling is plagued by similar guard-banding techniques. To cover the inaccuracy of the conventional parasitic models, designers are asked to use large design margins which prevent them from taking full advantage of an advanced process. Also, the excess design margin increases the time required for timing-closure, increasing design time, over design and ultimately the cost of chip design.

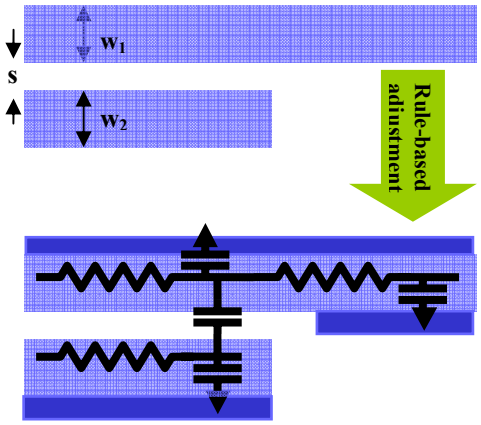

Figure 8: Rule-based adjustment of wire shapes attempt to model systematic shape variations.

As early as 2003, it was found that the shape dependence on interconnect line width, spacing and density had to be incorporated to remove the "unknown excess margin" in parasitic extraction accuracy [7]. A simple rule-based adjustment of wide widths and thicknesses based on spacing and density was introduced in RC extraction flows to account for shape variation as illustrated in Figure 8. However, in going to 90 and 65nm, the complexity of layout context influence on wire shapes is increasing to the point where these simplistic rule-based approaches are inadequate to predict systematic variations in interconnect delays "Now we are at the point where we simply need more complex models to reflect the expected shape of the interconnect" [8]. Just like OPC went from rule based to model based as optical effects became more complicated, design rules will need to follow in a similar direction.

IV. New Models for Sub-nanometer Variability

The increasingly non-linear behavior of interconnect and transistor shape variations requires a model based approach that can no longer be substituted by a set of rules. The impact of these variations must be injected into the current design flow to detect timing, noise and power failures that current methods are unable to capture.

To predict systematic shape variations, a model-based approach requires foundry manufacturing information to be captured in a predictive model that can be used at design time. This is graphically illustrated in Figure 9. These "compact models" must be built at the fabrication facility, and released to designers as SPICE models, DRC rules and parasitic extraction tables are today. This compact model, as described in [9], captures the entire manufacturing flow, including RET, OPC, lithography, etch, CMP and their interdependencies without disclosing manufacturing IP and provide critical information to designers and design tools to predict systematic manufacturing variations.

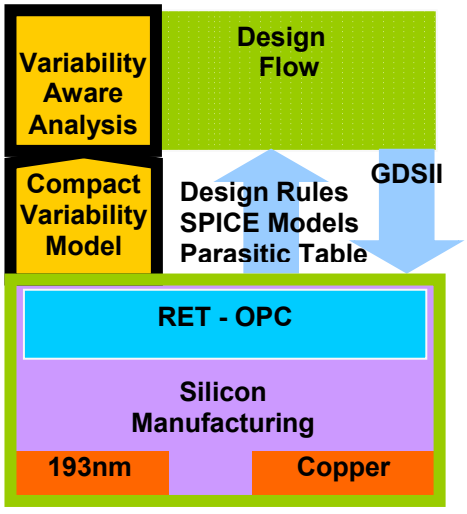

Figure 9: Compact model captures manufacturing flow bringing variability awareness to design flow.

It must have enough information to both predict systematic shape variations and their impact on interconnect parasitics and transistor behavior. The

impact of these variations must be seamlessly introduced into existing design tools and flows involving design rule checking (DRC), transistor parameter extraction, parasitic extraction (LPE), delay calculation, timing analysis, signal integrity and power analysis. These compact models may be encrypted to protect the intellectual property of foundries hidden in these models.

V. Timing Variability - Case Study

The impact of systematic on-chip shape variations on timing performance of a low-power 90nm chip using a compact variability model is shown in Figure 10 and Figure 11. Changes in transistor channel length and RC parasitics due to shape variation beyond what is captured in conventional SPICE and LPE models cause this variability in timing. Figure 10 and Figure 11 show the delay of each timing arc in the design vs. the number of arcs with that delay, both before ("ideal") and after ("actual") considering systematic shape variation. A systematic shift in the mean of the delays and a reduced variance can be seen.

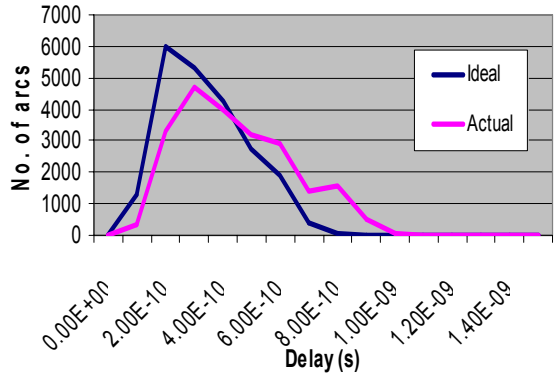

Figure 10: Systematic shape variations impact worst case timing.

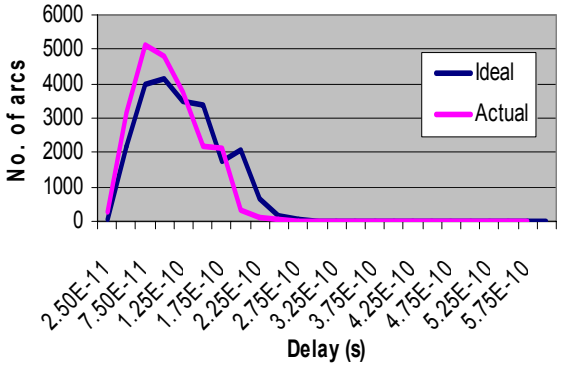

Figure 11: Systematic shape variations impact best case timing.

At worst case, Figure 10 shows that the systematic variation can slow down paths. Likewise, Figure 11 shows that at best case, the paths can get faster. The result is that setup and hold time violations can go undetected if a model for systematic shape variations is not included in timing analysis. Only a method that accounts for systematic variations is able to identify these parametric failures during design, reducing parametric yield failures.

VI. Conclusion

In this paper we described sub-nanometer manufacturing variations in IC design. The impact of these variations, if not addressed in the design, will cause significant manufacturing issues, such as poor yields, long yield ramp-up times and poor reliability. Realizing that DFM responsibility is shared between design and manufacturing; models and design methods must predict design sensitivity to these sub-nanometer effects during design. Once the sensitivities are known, a design engineer or EDA tool can optimize the design to reduce them. This results in shifting the mean and reducing the variance in the statistical distributions of key performance indicators, allowing performance requirements to be met.

VII. References

[1] J.P. de Gyvez and H.P. Tuinhout, "Threshold Voltage Mismatch and Intra-Die Leakage Current in Digital CMOS Circuits," *IEEE JSSC*, 2004.

[2] C. Malachowsky, (NVIDIA), *Advanced Reticle Symposium,* 2005

[3] C.H. Rahman and P.F. Zhang, "Three-dimensional DIBL for shallow-trench isolated MOSFETs," *IEEE Trans. Electron Devices*, 1999.

[4] P. Drennan, (Freescale Semi), WESCON 2005

[5] N.S. Nagaraj, "Benchmarks for Interconnect Parasitic Resistance and Capacitance," *Proc. ISQED,* 2003

[6] C. Bittlestone et. al., "Nanometer Design Effects and Modeling," *IEEE ISSCC Ckt Des For.*, 2005.

[7] NEC statement, 11th June, 2003

[8] Mark Bohr (Intel), "Chip interconnect draws designers, too", *EETimes*, June 6th 2005

[9] Philippe Hurat, Michel Cote, "A Genuine Design Manufacturability Checker for Integrated Circuit Designers", Bacus 2005