Lithography and Design in Partnership: A New Roadmap*

Andrew B. Kahng
UCSD Departments of CSE and ECE, La Jolla, CA 92093-0404 USA
abk@ucsd.edu

ABSTRACT

We discuss the notion of a 'shared technology roadmap' between lithography and design from several perspectives. First, we examine cultural gaps and other intrinsic barriers to a shared roadmap. Second, we discuss how lithography technology can change the design technology roadmap. Third, we discuss how design technology can change the lithography technology roadmap. We conclude with an example of the 'flavor' of technology roadmapping activity that can truly bridge lithography and design.

1. INTRODUCTION

In this paper, we explore the notion of a *shared* technology roadmap between lithography and design. The potential benefits of a 'shared roadmap' are immense: as noted in [8, 9], there is a promise of feasible (or, much less expensive) solutions to ITRS 'red bricks' (technology requirements with no known solution) [11], given appropriate new mechanisms for collaboration and synergy between disparate technology domains. However, it is difficult to envision, let alone attain, a shared roadmap between litho and design because (1) many cultural gaps separate the respective R&D communities, and (2) the driving metrics, and target values for these metrics, are not obvious. In this introductory section, we briefly illustrate the scope of existing cultural gaps; the concluding section of this paper illutrates how shared metrics and target values might be obtained in a principled manner for lithography and design.

Cultural gaps between lithography and design are apparent even at the level of basic concepts. These serve to alert us to the likelihood of 'disconnects'.

- To the lithographer, the design is simply a set of polygons; the origin of the polygons is unknown, and it is irrelevant whether the polygons comprise an arithmetic unit or a state machine or an A-to-D converter. To sell a wafer, CDs and WAT Ion/Ioff measurements must be verified. Going the other way, to the designer, lithography and indeed the entire manufacturing process is captured by the BSIM SPICE model and its corners; the origin and method of obtaining this model – and, in particular, its corners – are unknown. To sell a chip, chip-level power and timing specifications are verified on the tester.
- Different mindsets can also be seen in how we conceive of 'beyond the die'. The process or lithography engineer thinks of geometric variation and yield across wafers and lots. On the other hand, IC design engineers and tools think only inside the die: to a designer, 'beyond the die' means power and timing across package, board, and system.
- Finally, we can examine the ITRS itself, which contains explicit technology roadmaps for lithography and design. The Lithography roadmap has red and yellow bricks for metrics such as CD uniformity, data volume, linearity, and MEEF. By contrast, the Design roadmap has red and yellow bricks across such metrics as leakage power, mean time to failure, design reuse, and the number of circuit families that can be effectively combined on a single die using automated tools. It is challenging to identify sensible mappings between the respective sets of technology metrics and targets.

Does there exist a principled linkage between metrics and roadmaps of these two kinds of technologies? It is difficult to write down specific details, but such a linkage – if it exists – must be very broad, spanning the device architecture, the SPICE model, the key circuits (SRAM, logic, analog/RF), the product and, ultimately, cost. Given such a broad and amorphous linkage between lithography and design technologies, the key question is: Can we deconstruct this linkage so as to better align and drive the two technologies?

* This invited paper transcribes a plenary talk given on February 25, 2008 at the SPIE Advanced Lithography Symposium, San Jose, CA.

The remainder of this paper is organized as follows. In Section 2 below, we address the question of how lithography technology changes the design technology roadmap. Conversely, Section 3 examines how design technology can change the lithography technology roadmap. Finally, Section 4 addresses the need for, and potential benefits from, a principled sharing of the two roadmaps.

2. HOW LITHOGRAPHY CHANGES THE DESIGN ROADMAP

In this section, we suggest three examples of how Lithography technology changes the Design technology roadmap.

2.1. Restricted Design Rules (RDRs)

The exploding complexity of RETs, design rules, and physical effects such as pattern-dependent stress mean increased risk, margins, and cost. This trend is economically impossible to continue. As an industry, we must collectively understand that designers require freedom *from* choice; otherwise, it will be impossible to raise the level of abstraction and, correspondingly, the productivity of design. In this light, so-called 'restricted design rules' (RDRs) are inevitable.

Various regular design styles have been suggested to achieve reliable printability of subwavelength features. Gupta and Kahng [14] have pointed out that full-chip layouts may need to be assembled as a collection of regular printable patterns for technologies at 90nm and beyond. Lavin et al. [15] have proposed L3GO (Layout using Gridded Glyph Geometry Objects) with points, sticks and rectangle glyphs, to improve manufacturability. Using the glyph-based layout methodology, a circuit may avoid manufacturing challenges that arise from design irregularity. Liebmann et al. [16] have proposed a rule-based layout optimization methodology based on restrictive design rules (RDRs) to control linewidth on the poly layer. Having a limited number of linewidths along with single orientation of features, per RDRs, presents new challenges to automatic design migration. The work of Wang and Wong [17] studies the impact of grid-placed contacts on application-specific integrated circuit (ASIC) performance.

Looking forward to the 32nm and 22nm nodes, a grid-based layout scheme offers the promise of allowing layouts to be partitioned for double-exposure patterning [18]. Jhaveri et al. [19] discuss use of regular 'fabrics' that define the underlying silicon geometries of a circuit, and further discuss the benefits of constructing regular designs from a limited set of lithography-friendly patterns. Similarly, Maly et al. [20] proposed LDP (Lithographer's Dream Patterns), a methodology that incorporates extremely regular and uniform layout patterns with a large number of dummy patterns.

As noted in a recent work on SRAM bitcell design for 32nm interference-assisted lithography [32], there are several common justifications for slow adoption of RDRs. For example, (1) in logic design, the contact landing area and the diffusion patterns are problematic for a grid-based layout framework, while density remains the dominant competitive metric, particularly for consumer products; (2) in embedded SRAM design, static noise margin prevents uniform diffusion width; and (3) layout area is often observed to increase after migration to a regular layout framework. RDRs are inevitable nonetheless, and their arrival may be hastened by several key realizations. First, RDRs may actually bring area *benefits* rather than costs, e.g., if regularity can enable deployment of a *pushed rule*; [19] show this for an example D-flip-flop. Second, even if density worsens with RDRs, this is a 'one-time hit' or 'reset' from which the industry can straightforwardly scale; the benefits from easier and more timely node-to-node scaling transitions may far outweigh the costs of a one-time area increase. Finally, verdicts on RDRs likely have not correctly factored in the benefits that accrue from reduced margins and reduced turnaround times. We now discuss such a potential benefit that arises in the context of a 'reduced guardband' design methodology.

Reduced Guardband Implications for Design.

An important way in which lithography technology can change design outcomes is via the *model guardband* for design timing and power. When the timing guardband is tightened (e.g., better process control, single-orientation transistor gate layout, more restrictive pattern and pitch rules, etc.), the timing closure design task becomes easier and faster. Gates become smaller, hence area becomes smaller, hence wires become shorter – and there is a virtuous cycle of reduction. The net result, then, from reduction of model guardband is smaller die, and more die per wafer; automated design tools will run faster, as well. On the other hand, if the guardband is tightened but somehow without any process improvement, there will be more parametric yield loss during wafer sort - that is, more speed or leakage outliers will be scrapped. The industry also continuously weighs the economic viability of relaxing process variation limits in the technology roadmap [11].

Work at UCSD [1] gives the first-ever quantification of the impact of modeling guardband reduction on outcomes from the synthesis, place and route (SP&R) implementation flow. We assess the impact of model guardband reduction on various metrics of design cycle time and design quality, using open-source cores and production (ARM/TSMC) 90nm and 65nm technologies and libraries. Figures 1(a) and (b) show the impact of guardband reduction on area and routed wirelength, respectively, while Figure 1(c) shows the impact of guardband reduction on total SP&R flow runtime. We observe an average of 13% standard-cell area reduction and 12% routed wirelength reduction as the consequence of a 40% reduction in library model guardband.

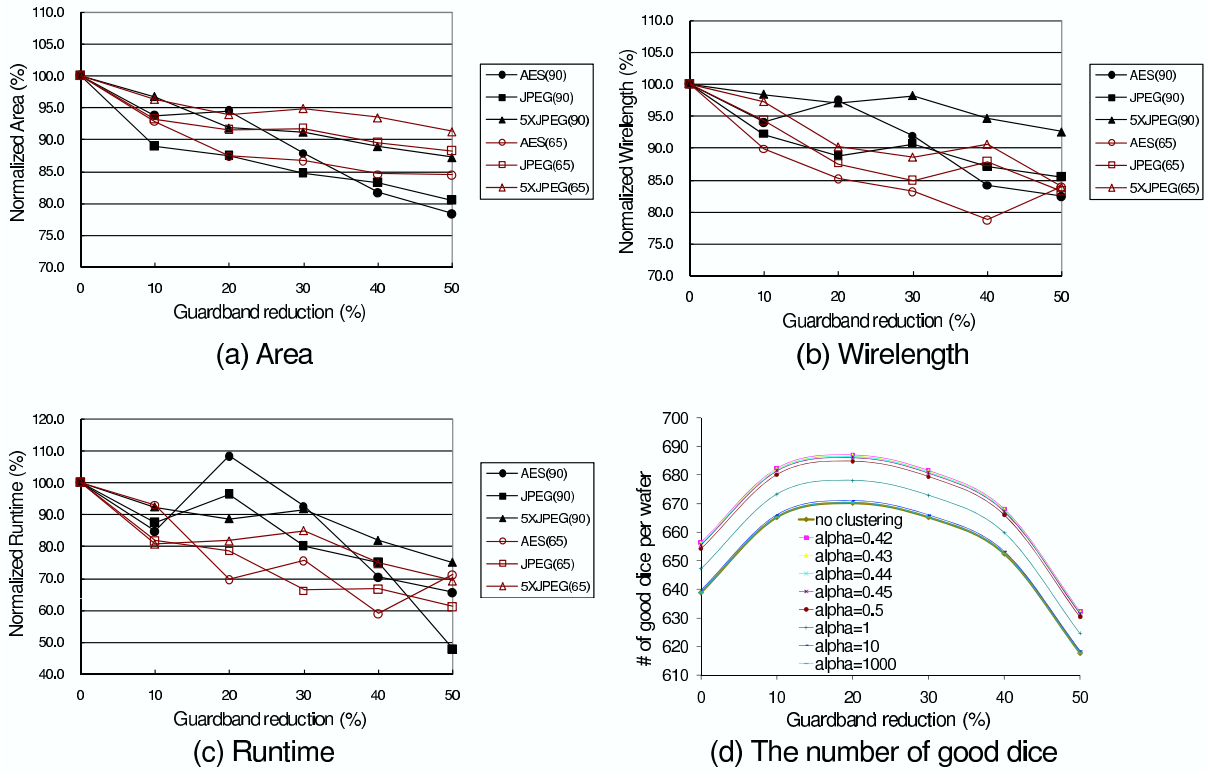

Figure 1: Impact of reduced guardband on design outcomes.

Of course, guardbanding exists in the process-to-design interface, and in the design methodology, to help guarantee high yield in the face of process variability. Overall yield may be modeled as the product of *random* yield, which depends on die area, and *systematic* yield, which is independent from die area: $Y = Y_r \cdot Y_s$. To assess the impact of guardband reduction on design yield, we track the change in the number of good dice per wafer as we reduce the design guardband. There are two scenarios: (1) we are able to control or improve the process so as to reduce the amount of guardbanding,[†] and (2) we simply apply a reduced guardband during the design process, even though the actual variability of the manufacturing process remains the same. Scenario (1) implies that Y_s remains at 3σ, while overall yield increases because we benefit from decreased random defect yield loss due to decreased die area. Scenario (2) changes Y_s and is more pessimistic because no process improvement is assumed: the design guardband reduction increases random defect yield Y_r due to reduced die area, but this trades off against decreased Y_s. Figure 1(d) plots the change in the number of good dice per wafer against the guardband reduction, for various defect clustering assumptions.[‡] We see a maximum increase of 4% in the number of good dice per wafer at approximately the 20% guardband reduction point. In that Moore's Law provides a rough equivalence of one week vs. 1%, the results of [1] suggest that design and process can together use model guardbanding as a lever to trade off time to market, die cost, and product area and power. Our results suggest that

[†]This Scenario (1) corresponds to performing 'iso-dense' (pitch- and Bossung curve-aware) timing analysis [2].

[‡]This plot reflects a typical SoC design in 90nm and 65nm, with die area = 1 cm^2 and only 50% of the die being logic that is affected by the guardband reduction (i.e., half of the die is memory, analog, etc.).

there is justification for the design, EDA and process communities to enable guardband recution as an economic incentive for manufacturing-friendly design practices.

2.2. Double-Patterning Lithography: Pattern Decomposition

A second example of how litho changes the design technology roadmap is in the context of double-patterning lithography (DPL). With DPL, two features must be assigned opposite colors (corresponding to different exposures) if their spacing is less than the *minimum coloring spacing* [21,23,25]. However, there exist pattern configurations for which pattern features separated by less than the minimum color spacing cannot be assigned different colors. In such cases, at least one feature must be *split* into two or more parts, whereby decreased pattern density in each exposure improves resolution and depth of focus (DOF). The pattern splitting increases manufacturing cost and complexity due to (1) generation of excessive line-ends, which causes yield loss due to overlay error in double-exposure, as well as line-end shortening under defocus; and (2) resulting requirements for tight overlay control, possibly beyond currently envisioned capabilities. Other risks include line edge (CD) errors due to overlay error, and interference mismatch between different masks. Therefore, a key optimization goal is to reduce the total cost of layout decomposition, considering the above-mentioned aspects. Recent work at UCSD [3] has developed an integer linear programming (ILP)-based solver for DPL pattern decomposition.

Figure 2, part (1), illustrates the ILP-based DPL coloring flow. Polygonal layout features in (a) are fractured into rectangles, over which the conflict graph is constructed and conflict cycles detected as shown in (b). To remove the conflict cycles in the conflict graph, a node splitting process is carried out along the dividing points denoted by the dashed lines in (b). The polygon splitting itself is governed by a process-aware cost function that avoids small jogging line-ends, maximizes overlap at dividing points of polygons, and preferentially makes splits at landing pads, junctions and long runs [23]. After the node splitting process for conflict cycle removal, the conflict graph is updated with newly generated nodes and updated edges as in (c). Finally, ILP-based coloring is carried out to obtain the final coloring solution on the 2-colorable conflict graph (d). An additional layout partitioning heuristic helps achieve scalability for large layouts.

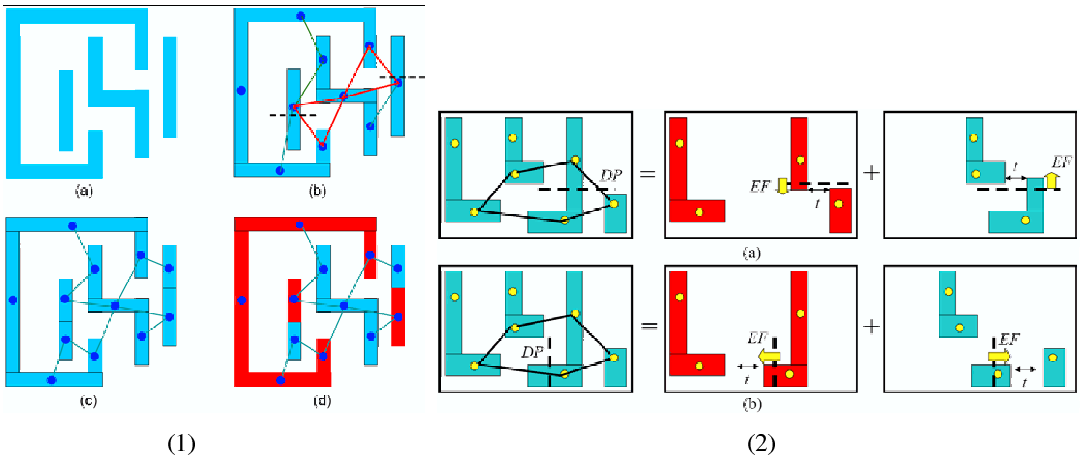

Figure 2: Overall DPL layout decomposition flow, and two examples of dividing points.

Splitting of a feature to meet DPL mask coloring requirements can be susceptible to pinching under worst process conditions of defocus, exposure dose variation and misalignment. Thus, two line-ends at a dividing point (*DP*) must be sufficiently overlapped. Figure 2, part (2), illustrates the extended features (*EF*) that are needed to address the overlap requirement, but that must also satisfy DPL design rules - i.e., the spacing between patterns at the dividing point must be greater than the minimum coloring spacing. The figure shows two layout decompositions that each remove the conflict cycle shown. However, when extending patterns for overlay margin, the dividing point in (a) causes a violation of the minimum coloring spacing, t, while the dividing point in (b) maintains the minimum coloring spacing even with line-end extension for overlay margin. [3] reports assessments of this approach using 45nm testcases. The ILP framework, being quite general, permits future enhancements of variability awareness via, e.g., (1) minimizing the difference between the pitch distributions of two masks, and (2) minimizing the number of distinct DPL layout solutions across all instances of a given master cell (to reduce variability between instances).

2.3. Double-Patterning Lithography: The Bimodal Analysis Challenge

A third example of litho changing the design roadmap is again related to DPL. When features are printed by two independent exposures, the 'odd' gates are printed with one mask, and the 'even' gates are printed with the other mask. As noted in [4], this results in a bimodal CD distribution, as well as a loss of both spatial correlation and line-to-space correlation. This can have far-reaching effects on design, where today's timing characterization and analysis methodologies implicitly assume a unimodal distribution, for which we take worst and best corners. The recent work of [4] experimentally studied the pessimism inherent in capturing a bimodal distribution with a unimodal framework.[§]

Figure 3(a) shows a bimodal CD distribution for 32nm technology measured from 24 wafers processed by DPL, as reported in [24]. We refer to the different CD distributions as corresponding to the different *colorings* (i.e., mask exposures) of the gate polys in a cell layout. Every cell instance in a design can be colored differently according to its location and the surrounding cell instances as shown in Figure 3(b). Therefore, instances of the same master cell in a timing path can be differently colored, and can have different electrical behaviors.

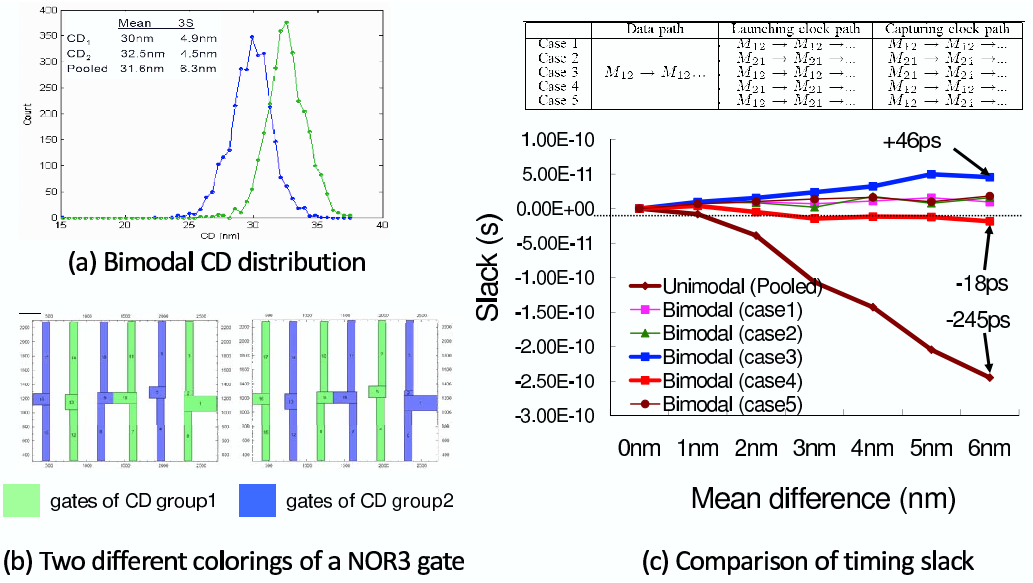

Figure 3: Bimodal CD distribution.

To see more explicitly and realistically the impact of bimodal CD distribution on the timing slack, we consider a most-critical timing path from an *aes* core, obtained as RTL from the open-source site *opencores.org* [12], which synthesizes to 40K instances, and is then placed and routed with a reduced set of 45nm library cells. Figure 3(c) shows the slack changes of each different coloring sequence of clock paths versus the mean difference of the two CD groups at the worst CD corner combination. In the figure, M_{12} and M_{21} refer to the two possible mask colorings (as in Figure 3(b)) of buffers in the clock distribution network. The timing path originally has zero slack when the CD mean difference is zero (i.e., two color groups have same CD mean). However, we observe that in the bimodal case4, where launching and capturing clock paths are colored by different sequences, the slack becomes negative, and a -18*ps* timing violation occurs at 6nm mean CD difference. At the same time, Figure 3(c) shows that the simplistic modeling of bimodal as unimodal results in unnecessarily pessimistic setup timing slack values. These data strongly suggest that the bimodal CD distribution arising in DPL must be dealt with accurately in both analysis and optimization. More generally, the industry must quickly comprehend whether a cell timing model is needed for each mask solution, whether a bimodal SPICE best-case / worst-case model is required, whether the design flow must perform incremental mask layout whenever the placement changes, etc. There is no question that DPL significantly affects the design technology roadmap.

[§]Even if a bimodal CD distribution is explicitly modeled, DPL will likely make timing closure more difficult for designers. While delay variation actually decreases when a timing path is made up of gates taken from two uncorrelated distributions, this is overshadowed when the timing analysis tool can no longer exploit spatial correlation to reduce worst-case pessimism.

3. HOW DESIGN TECHNOLOGY CHANGES THE LITHOGRAPHY TECHNOLOGY ROADMAP

We now discuss how Design technology can change the Lithography technology roadmap. Again, three examples are suggested.

3.1. The Power-Limited Frequency Roadmap for MPU Products

My research group at UCSD has been responsible for the ITRS clock frequency model updates in both the 2001 and the 2007 ITRS revisions. Clock frequency is a high-level way in which design changes the lithography roadmap. Specifically, in 2007, the ITRS finally acknowledged hard power constraints for microprocessors. The power limits, in turn, constrain maximum clock frequency and push the design itself in various ways (notably, to multicore architectures). As a result, the ITRS now projects only a 25% increase in processor clock frequency, per node.

The change in design approaches (lower-frequency multiprocessor architectures replacing higher-frequency uniprocessor architectures) affects lithography technology via a ripple of implications: less frequency means that the device (transistor) CV/I delay does not need to reduce as aggressively, and thus physical gate length does not need to decrease so deeply or rapidly into (mythical) red-brick values. The end result is that a slower frequency roadmap on the design side opens the door to a relaxed CD control requirement roadmap on the lithography (and FEP) side.

3.2. Impact of Design-Awareness

Design can also change lithography requirements by transmitting *design-awareness* to the manufacturing tools and process. Going forward, manufacturing can no longer think of 'design intent' as simply polygon shapes: design intent must encompass functionality and performance as well. As one example, timing slack on a signal path can be converted into power savings by reducing the transistor gate widths, or increasing transistor gate lengths, in the logic cells on the path. Alternatively, RET and litho accuracy requirements can be relaxed when there is timing slack. This is the basic idea of the 'MinCorr' Minimum Correction methodology developed at UCSD in 2002 [5], which has since been deployed in production design flows to perform transistor gate CD biasing for leakage power reduction.

Design awareness may also be exploited with respect to *redundancy*. Whether dummy fill for CMP uniformity, or redundant vias and wires for reliability, or spare cell instances to support metal-only design respins - certain shapes in the layout database do not require the standard level of exactitude in manufacturing. Understanding the growing use of design redundancy can change the scaling of RET and mask costs, not only through fracture count, but also through the inspection and defect disposition flows. Furthermore, if performance-critical structures are implemented with redundancy, then again it may be possible to relax lithography variability requirements.

3.3. Design for Equipment

A third way in which design can impact the process technology roadmap is via *Design For Equipment* (DFE). The common theme across potential DFE techniques is to make designs more variation-robust, and to thereby ease process control requirements. An early example [33] showed how lens aberration-aware placement can improve timing yield. Other possibilities abound - e.g., x- and y-direction overlay error budgets can potentially be traded off to maximize timing yield. In this subsection, we review recent work at UCSD [6] that exploits ASML DoseMapper technology to improve timing yield and leakage power.

DoseMapper [34, 35] allows for optimization of ACLV (Across-Chip Linewidth Variation) and AWLV (Across-Wafer Linewidth Variation) using exposure dose modulation. Figure 4(a) [37, 38] shows how slit exposure correction is performed by Unicom XL; the actuator is a variable-profile gray filter inserted in the light path. Overall, a correction range of $\pm$5% can be obtained with Unicom-XL for the full field size of 26mm in the X-direction. Scan exposure correction is realized by means of Dosicom, which changes the dose profile along the scan direction. The dose generally varies only gradually during scanning, but the dose profile can contain higher-order corrections depending on the exposure settings. The correction range for the scan direction is $\pm$5% (10% full range) from the nominal energy of the laser. When the requested X-slit and Y-scan profiles are sent to the lithography system, they are converted to system actuator settings (one Unicom-XL shift for all fields, and a dose offset and pulse energy profile per field). *Dose sensitivity* is the relation between dose and critical dimension, measured as CD [nm] per percentage [%] change in dose. Increasing dose decreases CD as illustrated in Figure 4(b), i.e., the dose sensitivity has negative value. A typical dose sensitivity D_s at $\leq$ 90nm is -2nm/% [36].

Today, the DoseMapper technique is used solely (albeit very effectively - e.g., [36]) to reduce ACLV or AWLV metrics for a given integrated circuit during the manufacturing process. However, to achieve optimum device performance (e.g., clock frequency) or parametric yield (e.g., total chip leakage power), not all transistor gate CD values should necessarily be the same. For devices on setup timing-critical paths in a given design, a larger than nominal dose (causing a smaller than nominal gate CD) is desirable, since this creates a faster-switching transistor. On the other hand, for devices that are on hold timing-critical paths, or in general that are not setup-critical, a smaller than nominal dose (causing a larger than nominal gate CD) will be desirable, since this creates a less leaky (although slower-switching) transistor. What has been missing, up to now, is any connection of such 'design awareness' – that is, the knowledge of which transistors in the integrated-circuit product are setup or hold timing-critical – with the calculation of the DoseMapper solution.

The work of [6] uses dose (computed by a quadratic programming optimization) to modulate gate poly CD across the exposure field, so as to optimize a function of delay and leakage power of the circuit. A complementary dose map-aware placement optimization considers systematic CD changes at different areas within a given dose map, and seeks to optimize circuit timing yield by selectively relocating timing-critical standard-cell instances. Initial simulation results reported in [6] show that more than 8% improvement in minimum cycle time of the circuit can be obtained at no cost of leakage power increase.

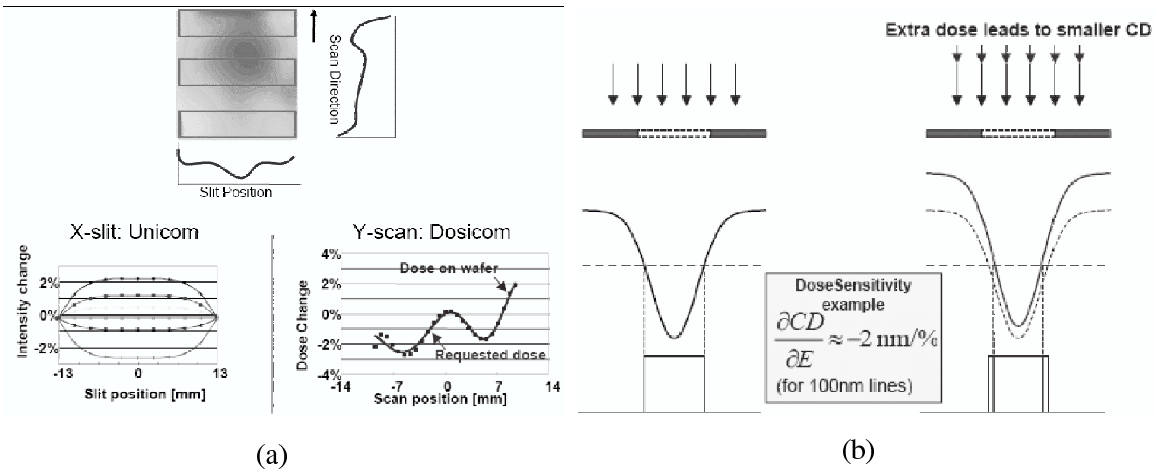

Figure 4. Dosemapper fundamentals. (a) Unicom-XL and Dosicom, which change dose profiles in slit- and scan-directions, respectively. Source: [37]. (b) Dose sensitivity: increasing dose (red color) decreases the CD. Source: ASML.

4. TOWARD A SHARED LITHOGRAPHY-DESIGN ROADMAP

Process *changes* what is possible, while Design *realizes* what is possible. When the semiconductor roadmap blithely projects the continued geometric increase of value - in a 'More Than Moore' sense - one may think that Process and Design, together, are successfully 'making it all happen'. Unfortunately, such thinking is wrong.

Today, Process and Design – or, Lithography and Design – are making life unnecessarily hard for each other. Designers are not using RDRs, process engineers are not listening to design intent, and both technology domains are rapidly becoming more expensive, unpredictable, and risky. The unfortunate reality is that More-Than-Moore electronic product value can stem from many other directions, e.g., embedded software, multi-core architecture, 3D integration, etc. In other words, workarounds are available for (lithography and/or design) technologies that become too risky and expensive. Eight years ago, in a keynote at the Japan DA Show [10], I called this the Dark Future for semiconductor process and design. An appropriate quotation, perhaps, is from Benjamin Franklin: "If we do not hang together, we will surely hang separately". With this in mind, the remainder of this section illustrates the 'flavor' of a more principled, and ultimately synergetic and beneficial, connection between the lithography and design technology roadmaps.

The Flavor of a 'Principled Connection': Line-End Taper Shape

We began by contrasting the ITRS metrics for lithography and design. In an ideal world, technologists would be able to take (1) lithography and RET metrics, and (2) electrical and design metrics, and then (3) from these, determine layout

practices and design rules. Various forms of utopian goodness might then follow, such as a principled division of technology requirements and R&D investments, and sharing of the burden of solving the red bricks in the roadmap. Here we illustrate this principled connection, using the example of *line-end tapering*.

Traditionally, lithography and OPC engineers seek to make the line-end as rectangular as possible, while meeting line-end gap (or, LEG) and line-width at gate edge (or, LW0) requirements. Though these geometric metrics have served as good indicators, ever-rising contribution of line-end extension to layout area necessitates reducing pessimism in qualifying line-end patterning. The quality of line-end patterning depends on rounded shape of line-end, as well as line-end gap and linewidth at device. We use the word *'taper'* to describe the shape of a line-end. Line-end itself is defined as the extension of polysilicon shape beyond the active edge.

Work reported in [7] employs a 3D TCAD simulator [13] to investigate device capacitance, I_{on} and I_{off} according to various taper shapes and line-end extension lengths. In these preliminary results, Ion can change by as much as 4.5% and $Ioff$ by as much as 30%, as shown in Figure 5(b).

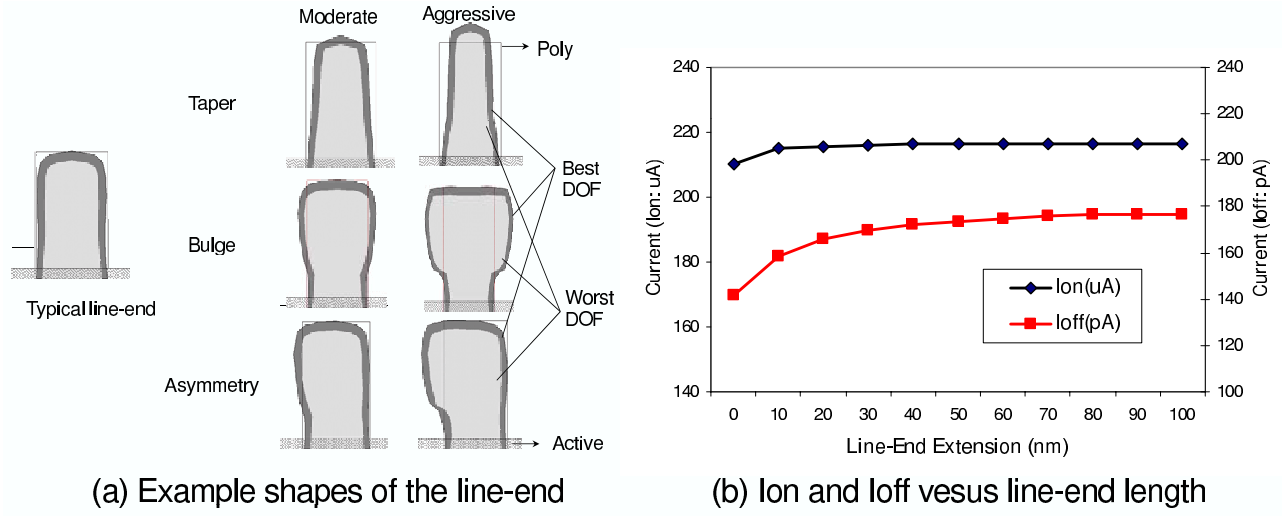

Figure 5: Example shapes of the line-end (a), and change of I_{on} and I_{off} with different line-end lengths (b).

The work of [7] further develops a modeling framework that includes (1) parametric specification and generation of taper shapes, (2) capacitance modeling of the parameterized taper shapes, and (3) modeling of Ion and Ioff with nonuniform impact across the gate width from the capacitance model.

Taper Shape Generation. Relying on OPC and simulation severely complicates and limits flexibility in exploring taper shapes. Therefore, we automatically generate taper shapes according to a *super-ellipse* template. The super-ellipse curve is defined by the equation '$\left|\frac{x}{a}\right|^n + \left|\frac{y-k}{b}\right|^n = 1$' where $n > 0$ and a and b are the radii of the oval shape, and k represents taper shift in the y-axis. For a given taper shape, a and b represent gate length and length of taper, respectively; n determines the slope, or corner routing, of the taper.

Capacitance Modeling. Gate capacitance is a sum of channel capacitance ($C_{channel}$) and line-end extension capacitance (C_{taper}). C_{taper}, in turn, is a fringe capacitance between gate extension and the channel. Due to the large difference in electric field between the gate edge and the gate extension, we can simply model the capacitance of gate extension as a sum of fringe capacitance of each segment and the capacitance of the gate edge which is the fringe capacitance without line-end extension.

I_{on} and I_{off} Modeling. Using the capacitance model for taper, a model for taper impact on I_{on} can be developed. Gate segments near the gate edge are affected more by the taper capacitance, and the taper effect decreases exponentially as the distance from the gate edge increases. The i_{on} of an individual gate segment s is a function of capacitance of the line-end extension; I_{on} is the sum over all segments. In the following equation, s is the segment index that accounts for the distance

from the gate edge. C_{taper_top} and C_{taper_bottom} represent taper capacitances at top and bottom gate edges, respectively. Then, total I_{on} current may be expressed as

$$I_{on} = \sum_{s=1}^{N} i_{on}(C_{taper_top}, C_{taper_bottom}, s, L) \quad (1)$$

$$i_{on}(C_{taper_top}, C_{taper_bottom}, s, L) = i^{0}_{on}(L_s) + \Delta i_{on}(C_{taper_top}, s, L_s) + \Delta i_{on}(C_{taper_bottom}, N-s+1, L_s) \quad (2)$$

Here, $i^{0}_{on}(L_s)$ is the base current of the gate segment, as measured from a large-width device which is not affected by line-end extension. The additive current (Δi_{on}) for each segment of the gate is modeled as a function of the line-end capacitance (C_{taper}), segment index (s) and linewidth of the segment (L_s). I_{off} is modeled similarly, but has an exponential relationship with the gate length L; an exponential function is used to model the I_{off} change with gate length.

Misalignment Modeling. Misalignment changes the length and width of segments near the channel. Since segments inside the channel affect I_{on} and I_{off} differently than those inside the taper, it is necessary to determine whether a given segment belongs to channel or taper. Assuming perfect printing of diffusion,only Y-direction misalignment needs to be modeled. For a given misalignment condition, we may calculate the I_{on} and I_{off} of the entire gate by summing up segment currents (i_{on} and i_{off}). If the probability of misalignment m is $P(m)$, and the current under such misalignment is $I(m)$, then the expected current I_{exp} with misalignment modeling is calculated as $I_{exp} = \sum_{m=1}^{N_{sites}} P(m)I(m)$.

The framework reported in [7] enables evaluation of electrical metrics of line-end shapes, leading to principled derivation of less pessimistic line-end extension rules, and improved criteria for OPC correction of line-end shapes. For example, Figure 6(a) shows an example of bitcell layout, and Figure 6(b) shows the corresponding layout constraint graph that defines the width of the bitcell. Figure 6(c) shows the tradeoff curve under given design rules for 65nm technology. From Figure 6(c), we observe that if we permit a factor of 2 leakage increase, we can reduce the line-end extension design rule to about $50nm \sim 20nm$, and reduce the bitcell size by about $7.69\% \sim 12.31\%$, depending on the super-ellipse exponent which represents the cost or complexity of OPC and lithography.

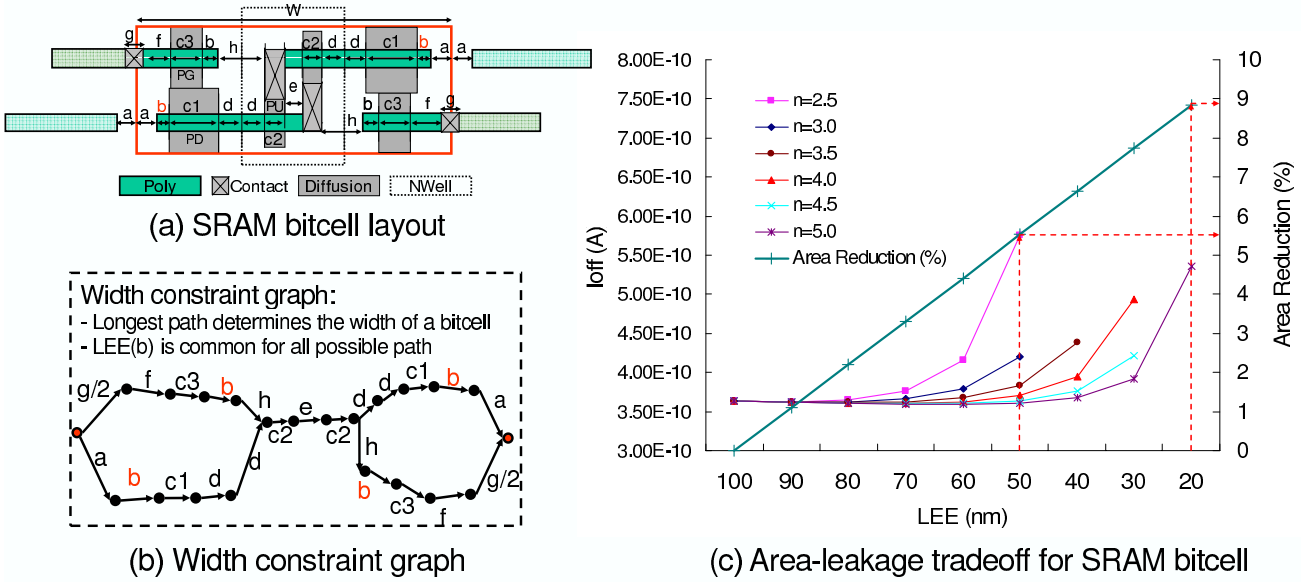

Figure 6: Area and leakage tradeoff for an SRAM bitcell with respect to line-end length and tapering.

5. ACKNOWLEDGMENTS

Many thanks are due to Puneet Gupta, Chul-Hong Park, Kwangok Jeong, Kambiz Samadi and Hailong Yao for their collaborations on the research at UCSD cited above, and for their help in compiling both the SPIE Advanced Lithography presentation and this transcription.

REFERENCES

1. K. Jeong, A. B. Kahng and K. Samadi, "Quantified Impacts of Guardband Reduction on Design Process Outcomes", *Proc. ISQED*, 2008, pp. 790–897.
2. P. Gupta and F.-L. Heng, "Toward a Systematic-Variation Aware Timing Methodology", *P*roc. DAC, 2004, pp. 321–326.
3. A. B. Kahng, C. -H. Park, X. Xu and H. Yao, "Layout Decomposition for Double Patterning Lithography", *Proc. ICCAD*, 2008, to appear.
4. K. Jeong and A. B. Kahng, "Timing Analysis and Optimization Implications of Bimodal CD Distribution in Double Patterning Lithography", *Proc. ASPDAC*, 2009, to appear.
5. P. Gupta, A. B. Kahng, P. Sharma and D. Sylvester, "Method for Correcting a Mask Layout", U.S. Patent No. 7149999, Dec. 2006.
6. K. Jeong, A. B. Kahng, C. -H. Park and H. Yao, "Dose Map and Placement Co-Optimization for Timing Yield Enhancement and Leakage Power Reduction", *Proc. DAC*, 2008, pp. 516–521.
7. P. Gupta, K. Jeong, A. B. Kahng and C. -H. Park, "Electrical Metrics for Lithographic Line-End Tapering", *Proc. PMJ*, 2008, pp. 70238A-1–70238A-12.
8. A. B. Kahng, "The Road Ahead: Shared Red Bricks", *IEEE Design and Test*, 2002, pp. 70–71.
9. A. B. Kahng, "The Road Ahead: How Much Variability Can Designers Tolerate?", *IEEE Design and Test*, 2003, pp. 96–97.
10. A. B. Kahng, "Futures for DSM Physical Implementation: Where is the Value, and Who Will Pay?", *keynote address*, 12th Japan DA Show, Tokyo, July 12, 2000.
11. *International Technology Roadmap for Semiconductors*, http://public.itrs.net/.
12. *OPENCORES.ORG*, http://www.opencores.org/.
13. Davinci User's Guide, Version Y-2006.06-SP1.
14. P. Gupta and A. B. Kahng, "Manufacturing-Aware Physical Design", *Proc. ICCAD*, 2003, pp. 681–687.
15. M. Lavin, F.-L. Heng and G. Northrop, "Backend CAD Flows for Restrictive Design Rules", *Proc. ICCAD*, 2004, pp. 739–746.
16. L. Liebmann, D. Maynard, K. McCullen, N. Seong, Ed. Buturla, M. Lavin and J. Hibbeler, "Integrating DfM Components Into a Cohesive Design-To-Silicon Solution", *Proc. SPIE DPIMM*, 5756 (2005), pp. 1–12.
17. J. Wang and A. K. Wong, "Effects of Grid-Placed Contacts on Circuit Performance", *Proc. SPIE CPICC*, 5043 (2003), pp. 134–141.
18. J. Wang, A. K. Wong and E. Y. Lama, "Standard Cell Design with Regularly-Placed Contacts and Gates", *Proc. SPIE DPIMM*, 5379 (2005), pp. 56–66.
19. T. Jhaveri, L. Pileggi, V. Rovner and A. J. Strojwas, "Maximization of Layout Printability/Manufacturability by Extreme Layout Regularity", *Proc. SPIE DPIMM*, 6156 (2006), pp. 67–81.
20. W. Maly, Y.-W. Lin and M. Marek-Sadowska, "OPC-Free and Minimally Irregular IC Design Style", *Proc. DAC*, 2007, pp. 954–957.
21. G. E. Bailey et al., "Double Pattern EDA Solutions for 32nm HP and Beyond", *Proc. SPIE DPIMM*, 2007, pp. 65211K-1–65211K-12.
22. C. Chiang, A. B. Kahng, S. Sinha, X. Xu, and A. Zelikovsky, "Bright-Field AAPSM Conflict Detection and Correction", *Proc. DATE*, 2005, pp. 908–913.
23. M. Drapeau, V. Wiaux, E. Hendrickx, S. Verhaegen and T. Machida, "Double Patterning Design Split Implementation and Validation for the 32nm Node", *Proc. SPIE DPIMM*, 2007, 652109-1–652109-15.
24. M. Dusa et al., "Pitch Doubling Through Dual-Patterning Lithography Challenges in Integration and Litho Budgets", *Proc. SPIE OM*, 2007, pp. 65200G-1–65200G-10.
25. J. Finders, M. Dusa and S. Hsu, "Double Patterning Lithography: The Bridge Between Low k1 ArF and EUV", *Microlithography World*, 2008.
26. A. B. Kahng, X. Xu and A. Zelikovsky, "Fast Yield-Driven Fracture for Variable Shaped-Beam Mask Writing", *Proc. SPIE PMJ*, 2006, pp. 62832R-1–62832R-9.
27. C. Lim et al., "Positive and Negative Tone Double Patterning Lithography for 50nm Flash Memory", *Proc. SPIE OM*, 2006, pp. 615410-1–615410-8.
28. C. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*, Wiley, 2007.
29. M. Maenhoudt, J. Versluijs, H. Struyf, J. Van Olmen, and M. Van Hove, "Double Patterning Scheme for Sub-0.25 k1 Single Damascene Structures at NA=0.75, λ=193nm", *Proc. SPIE OM*, 2005, pp. 1508–1518.
30. W.-Y. Jung et al., "Patterning With Spacer for Expanding the Resolution Limit of Current Lithography Tool", *Proc. SPIE DPIMM*, 6125 (2006), pp. 61561J-1–61561J-9.
31. J. Rubinstein and A. R. Neureuther, "Post-Decomposition Assessment of Double Patterning Layout", *Proc. SPIE OM*, 2008, pp. 69240O-1–69240O-12.
32. R. T Greenway, K. Jeong, A. B. Kahng, C.-H. Park and J. S. Petersen, "32nm 1-D Regular Pitch SRAM Bitcell Design for Interference-Assisted Lithography", *Proc. SPIE PT*, 2008, to appear.
33. A. B. Kahng, C.-H. Park, P. Sharma and Q. Wang, "Lens Aberration-Aware Timing-Driven Placement", *Proc. DATE*, 2006, pp. 890–895.
34. G. Zhang, M. Terry, S. O'Brien *et al.*, "65nm node gate pattern using attenuated phase shift mask with off-axis illumination and sub-resolution assist features", *Proc. SPIE OM*, 5754 (2005), pp. 760–772.
35. N. Jeewakhan, N. Shamma, S.-J. Choi *et al.*, "Application of dosemapper for 65-nm gate CD control: strategies and results", *Proc. SPIE PMJ*, 6349 (2006), pp. 63490G.1–63490G.11.
36. J. B. van Schoot, O. Noordman, P. Vanoppen *et al.*, "CD uniformity improvement by active scanner corrections", *Proc. SPIE OM*, 4691 (2002), pp. 304–314.
37. http://wps2a.semi.org/cms/groups/public/documents/membersonly/van_schoot_presentation.pdf.
38. http://www.asml.com/