Litho/Design Co-optimization and Area Scaling for the 22-nm Logic Node

James W. Blatchford[a], Steven L. Prins[a], Scott W. Jessen[a], Thuc Dam[b], KiHo Baik[b],
Linyong Pang[b] and Bob Gleason[b]

[a]Advanced CMOS Technology Development, Texas Instruments, Inc., Dallas, TX, USA
[b]Luminescent Technologies, Palo Alto, CA, USA

We present a comprehensive study of area scaling for 22nm-logic-node routed metal/via layers as a function of route pitch and patterning strategy in both single-exposure (SE) and double-patterning (DP) regimes. For each candidate route pitch (88-56nm), we determine an optimal illumination scheme and develop layout rules for the metal layers. A perturbative area model is used to approximate the impact of the candidate rule set on area scaling. For the most promising SE case, we apply a novel 'source/design optimization' technique to further optimize illumination and rules, wherein we extend the source-mask optimization approach (1) by allowing design rules to vary in the analysis. We demonstrate that the optimal area scaling achievable with DP techniques can be vastly superior to SE, and therefore may justify the associated additional cost per wafer.

Motivation

The 22-nm logic node and its associated 20-nm 'half-node' are unprecedented in that both must be realized without the benefit of an enabling lithographic exposure tool. As a result, many device manufacturers are investigating double-patterning (DP) techniques to enable device area scaling with current 1.35NA 193nm exposure tools. The practical pitch resolution limit for line-and-space layers using single-exposure (SE) at 1.35NA is generally accepted to be 76nm (2). DP techniques show great promise for extending pitch resolution well below this limit (3); however, the added cost per wafer for these techniques can be very high depending on the number of critical layers for which they are implemented. Of particular interest are the back-end-of-line (BEOL) layers, typically MET2 through MET6 in a low-power CMOS design. For these layers, determining whether SE or DP is optimal for overall die cost requires detailed analysis of the tradeoff between added cost per wafer and the smallest die area achievable with each technique. Complicating matters is the fact that the resolution enhancement techniques necessary to achieve tight metal route pitch—e.g., dipole illumination (2)—typically have an adverse effect on ancillary design rules such as minimum metal-to-metal spacing at line ends.

In this work, we summarize a detailed study of the optimal BEOL area that can be achieved for both SE and DP techniques as a function of metal route pitch. We leverage the fact that these layers are auto-routed to develop a relatively simple model for area scaling as a function of a handful of layout rules. At each value of route pitch (88-56nm), we optimize illumination and patterning schemes and develop a full set of associated design rules (DRs). For the most promising SE case, we further optimize illumination and rules using a new technique, 'source-design optimization', where we extend a traditional source-mask optimization approach (1) by allowing key DRs to vary in the analysis. We demonstrate that the optimal area scaling achievable with DP techniques is vastly superior to SE, therefore justifying the concomitant additional cost per wafer.

Area Model and Rule Definitions

To assess the area impact of rule choices, we constructed an approximate area model derived from a 50x50um densely routed layout clip from a 45nm low-power CMOS part based on an 8-track standard logic library. Only layers MET2 and above were considered. From this clip, we made several observations: First, all routing is almost exclusively at minimum metal width; i.e., larger widths can be neglected in the area calculation. Secondly, assuming the main routing direction is horizontal, the number of cross-connects (the vertical segment in Fig. 1a) is relatively small, and connections spanning more than three metal tracks are extremely rare. Thirdly, line-ends occur exclusively in the horizontal direction. As a result, the layout can be described in a compact set of rules, illustrated in Fig. 1a. Here d1 represents the line-end-to-cross-connect (LE-CC) space, d2 the cross-connect (CC) width, d3 the horizontal metal (‘wire’) length, and d4 the line-end-to-line-end (LE-LE) space. Note that for the 45nm clip as drawn, d1=d4=110nm, d2=70nm, d3=220nm, and route pitch is 140nm. The evaluation of area for each rule combination is performed using a perturbative algorithm, the details of which are too extensive to review in this short document. In a nutshell, for each candidate route pitch and rule set, we first ‘scale up’ the rule values by the ratio of the candidate route pitch to that of the 45nm clip (140nm). Starting at MET2, we then attempt to ‘rewire’ this clip with the scaled-up candidate rule set assuming the positions of all vias and cross-connects are fixed, and that the topology of the layout must be maintained. When conflicts occur in, e.g., LE-LE spacing, an entire row of routed cells is shifted by a unit of area equal to gate minimum contacted pitch x route pitch x the number of tracks in the cell (in this case, 8). Once the entire clip has been rewired in this fashion, we add up the additional area to assess the die scaling impact, expressed as a percent of the total clip area. A similar analysis is performed at MET3. Contributions from increasing rule values for MET4 and above (aside from route pitch) were found to be very small and were therefore neglected.

Figure 1b shows the sensitivity of die area to increases in the value of each rule, as predicted by this perturbative area model. Here the rule values have been expressed as a multiple of their ‘fully scaled’ value, calculated by taking the corresponding 45nm logic rule and ‘scaling it down’ by the ratio of the candidate route pitch to that of the 45nm clip.

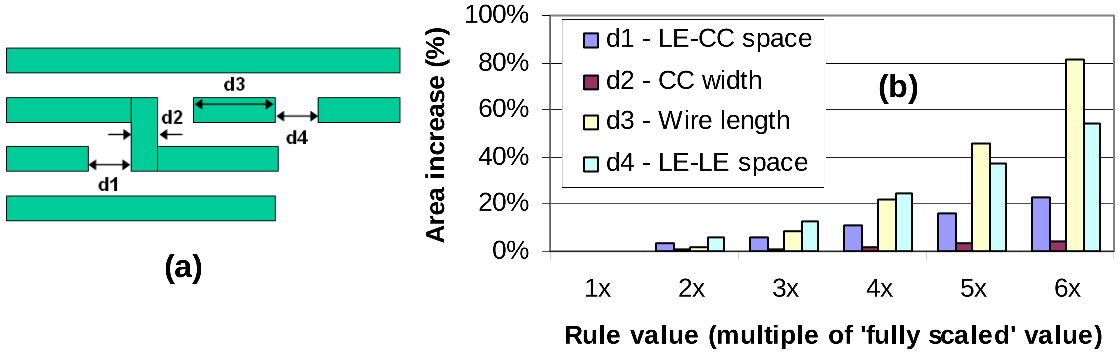

Figure 1. (a) Schematic illustrating design rules describing the reference 45nm logic clip for device area calculations. (b) Sensitivity of routed area to increases in individual rule values, expressed as multiples of their nominal value in the clip.

In other words, for a given route pitch, the 'fully scaled' value for a given rule is the value for which the candidate layout can be linearly shrunk to that route pitch without an area increase. From Fig. 1b, it is evident that substantial area increases only occur as a rule reaches ~3x its fully scaled value. In this regime, the LE-LE space rule (d4) has the largest area impact, followed by the wire length (d3) and LE-CC space (d1). Note that area is a relatively insensitive function of the CC width (d2).

Analysis: Rules and Patterning Strategies

We now examine optimal rules, illumination and patterning schemes as a function of route pitch. For this study, we considered route pitches from 88nm to 56nm in 4-nm increments. At each route pitch, an optimized illuminator was constructed with a 'source mapping' technique (4) using design clips representative of each rule, with image log-slope in photoresist as the optimization variable. We then used this illuminator to determine design rules d1-d4 based on a larger number of clips chosen by considering all possible combinations of structures with cross-connects 0-3 tracks high, with these requirements for each clip: Depth of Focus (DoF) >100nm; Mask Error Enhancement Factor (MEEF) <3.5 for main feature, <7 for line ends. A commercial lithography simulator was used with an internal 28nm-class lumped-parameter model for the evaluation of litho metrics. In the analysis, the final metal target was assumed to be half of the route pitch (e.g., 32nm at 64nm pitch); however, we allowed the exposure bias (i.e., mask CD) and litho CD target to vary globally to achieve the best MEEF and DoF response. In cases where the optimum litho target was larger than the final Si target, we assumed an etch process could be tailored to reduce the CD to the target value, and that etch biasing produced a CD change at a line-end 1.5x larger than that of a long metal line.

Figure 2 shows the evolution of the rules, patterning strategy and illumination scheme as route pitch is decreased. Again, rule values are expressed as multiples of their fully scaled values, per the discussion above. For the 88 and 84nm route pitch cases, a

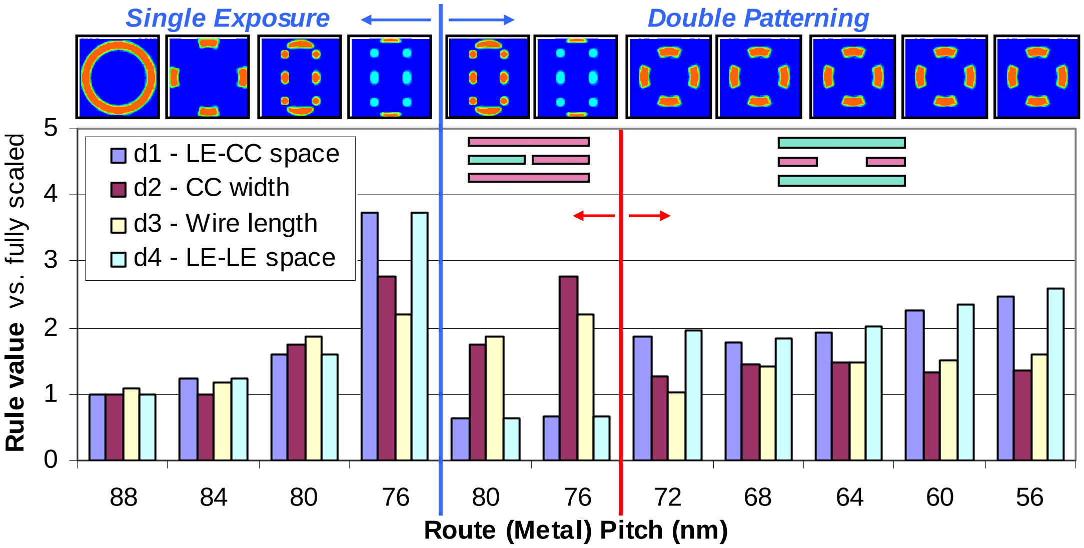

Figure 2. Simulated optimal illumination conditions and rule values as a function of metal route pitch. Results are segregated by patterning approach—single-exposure or double-pattern (DP)—as indicated at the top of the Figure, and the DP results by pattern-split and pitch-split DP, as indicated in the inset schematics.

straightforward annular or CQuad illuminator can be used, and the rules can be realized with single exposure with little deviation from full scaling. Pitches of 80 and 76nm also are realizable with SE; however, a strong dipole component along the Y-axis is required. At 80nm, other 'poles' in the illuminator can be present and optimized to give balanced performance for each rule; at 76nm any additional 'off axis' light leads to unacceptably large MEEF for dense features. The strong asymmetry of the illuminator leads to a substantial deviation from scaling for the design rules, especially in the 76nm-pitch case. For this pitch, rules d1 through d4 are required to be ~4x their scaled value to be imageable with the above minimum lithographic requirements.

Considering the substantial deviations from scaling in the 80 and 76nm pitch regimes, we also studied these pitches with DP techniques. Here, we found the best approach to be to re-use the optimal illuminator from the corresponding SE case, combined with a 'pattern split' DP approach. Figure 3 elucidates this technique schematically for two key cases. Figure 3a represents the drawn intent, and Fig. 3b shows optimal DP decomposition strategies using pattern-split DP. Since the illumination approach is capable of resolving the tight route pitch with a single exposure, features on adjacent route tracks can be freely moved to either mask to allow for optimum line-end spacing, as is demonstrated in the Figure. One drawback of this approach is that the CC width, d2, must be increased; however, this rule has minimal impact on area per the previous section. Returning to Fig. 2, we see the impact of this DP strategy in the 80 and 76 nm regime is to reduce the LE-LE and LE-CC spacing rules, d1 and d4, to a very small (hyperscaled) value, while the other two rules remain unchanged. We note this approach differs substantially from other reports in this regime that use less aggressive illumination and a 'stitching' method to resolve dense features adjacent to small metal gaps (5).

Finally, for pitches of 72nm and below, a 'pattern split' approach is impossible due to resolution limits, and a 'pitch split' DP strategy must be invoked, wherein the illumination is tailored to a pitch 2x larger than the desired route pitch, and all features along a given route track must be placed on the same mask. A single CQuad aperture is sufficient to cover this pitch regime; therefore, rules for each route pitch are very similar, to within slight variations based on etch bias. This DP strategy is illustrated schematically in Fig 3c. One implication of this approach is that adjacent line-ends in a dense environment are required to be on the same mask, driving the LE-LE spacing rule d4 to ~2x its fully scaled value. Furthermore, two-track cross-connects must be split between the two masks with a overlapped region between tracks that is nominally ~20nm but possibly much larger depending on etch bias. This results in a sub-resolution imaging environment in the region indicated by the circle in Fig 3c, leading to a very high MEEF value, optimally only ~6.7, for the feature adjacent to the cross-connect. We therefore chose to adopt an alternate decomposition strategy, as illustrated in Fig. 3d, and enforce a

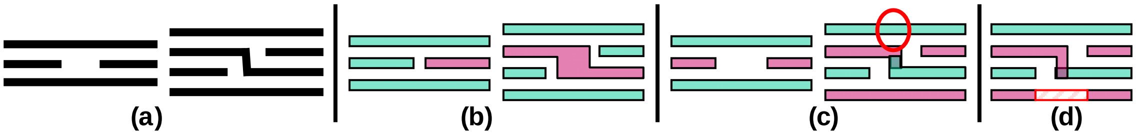

Figure 3. Comparison of 'pattern split' and 'pitch split' DP strategies. (a) Desired pattern. (b) Optimal result using pattern split DP at 76nm route pitch. (c) Optimal result using pitch split DP at 64nm route pitch. The red circle indicates a marginal area. (d) Alternate pitch split DP strategy involving a keep-out zone (cross-hatched area).

forbidden zone on one side of the cross-connect, indicated by the cross-hatched area in the Figure. We accounted for the impact of this restriction in the area calculations by adding a correction equal to the area of routing blocked, (2 · d1 + d2) · route pitch. We also investigated converting the cross-connect to a diagonal pattern under annular illumination (5); however, the best MEEF obtained for the adjacent feature after detailed optimization was still only ~5.5, unacceptably high compared to the target value of <3.5.

80nm Pitch with Single Exposure: Source/Design Optimization

As the 80nm pitch regime is the most promising for area scaling with SE, we performed further detailed optimization of illumination and rules in this regime using a new technique we call 'Source/Design Optimization' (SDO). Again, space limitations prevent detailed discussion of the approach. Starting from the candidate rule set and illuminator from the previous section, we develop a 'row' of structures representing limiting cases for each rule, as is indicated schematically in Fig. 4a. Expanding on the method of Chang *et al.* (7), we construct a large module that contains many copies of this row of structures, each copy generated with a slightly different (generally tighter) combination of the rules d1-d4. We then use source-mask optimization techniques (1) to develop an illuminator that allows as many rows as possible to pass the litho success criteria stated above, with each row individually weighted based on the area model above.

Figure 4b compares the illuminator developed in the previous section, OP1, with the optimal parameterized and free-form sources—OP2 and FF, respectively—derived from SDO. Figure 4c compares the design rules extracted from each illuminator and predicted area scaling impact. The OP1 illuminator provides the best balance of rule scaling; however, it has a fairly high associated area increase of ~9% vs. full scaling. OP2 and FF both give superior line-end performance (d1 and d4) at the expense of the CC width and wire length rules (d2 and d3). Both illuminators provide superior area performance to OP1, with OP2 giving an area increase of only ~3% over the fully scaled value.

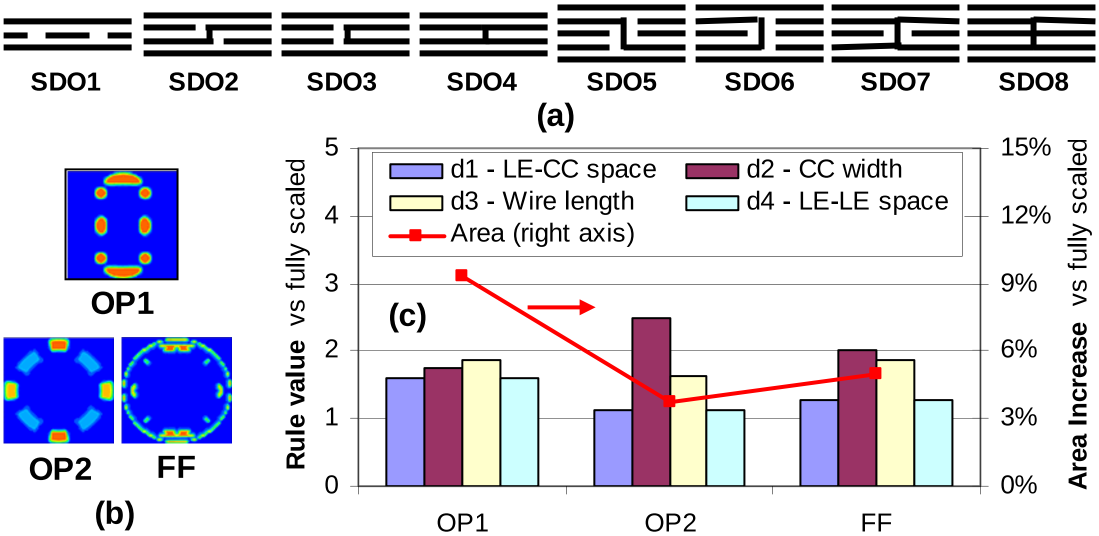

Figure 4. (a) Schematic representation of structures used for source/design optimization (SDO) experiments. (b) Initial illumination condition (OP1) and best parameterized (OP2) and free-form (FF) illuminators as predicted by SDO. (c) Comparison of design rules (left axis) and predicted die area increase (right axis) for the three illuminators in (b).

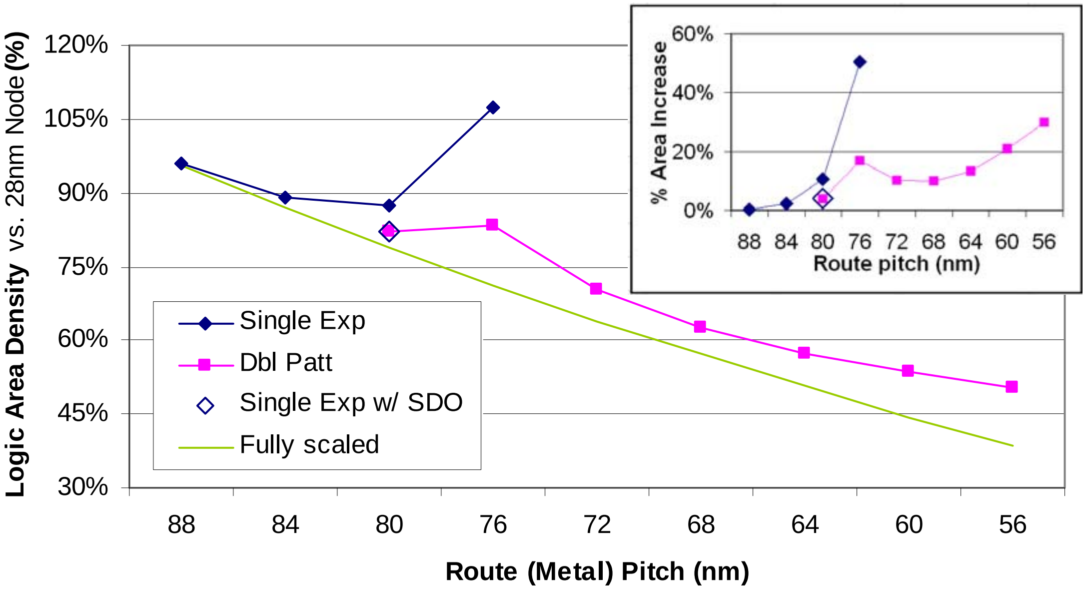

Figure 5. Logic area scaling for each design rule and patterning strategy presented in Fig. 2, as predicted by the area model, expressed as reduction vs. the same layout in 28nm. The open diamond represents the best result using SDO, per Fig. 4. The solid curve represents the 'fully scaled' result. Inset shows area increase above the fully scaled result.

Conclusions – Area Analysis

Figure 5 shows area scaling predictions in each regime based on the rules established in Fig. 2. Here, the solid curve represents the logic area achievable if all rules scaled equally from the 45nm clip. Note that the scaling numbers are expressed as an area reduction from the 28nm node (at 90nm route pitch), as node-on-node scaling is the most appropriate comparison for determining optimum strategy. The inset shows the increase in area over the fully scaled value. In the single-exposure regime, 80nm route pitch is clearly the best candidate, with the result using the SDO technique (open diamond) approaching the fully-scaled limit. In the double-patterning regime, area is found to decrease monotonically for all route pitches considered, in spite of an increasing deviation from the fully-scaled result as route pitch is reduced. The best achievable area with DP is ~2x better than that achievable with SE techniques. We therefore conclude that the optimal exposure strategy for 22nm logic BEOL layers is DP with a very tight route pitch, and that the substantial area improvement will outweigh the additional cost associated with this technique. We acknowledge that other criteria such as device performance with decreasing gate width may alter this conclusion.

References

1. See, for example, L. Pang *et al.*, Photomask Technology Proc. **7122** (2008).
2. A. Engelen *et al.***,** Proc. SPIE **7274**, 72741K (2009).
3. J. Finders *et al.*, Proc. SPIE **6924**, 692408 (2008).
4. See, for example, G. Zhang *et al.*, Proc. SPIE **5754**, 760 (2005).
5. V. Wiaux *et al.*, Proc. SPIE **6924**, 692409 (2008).
6. V. Wiaux, presented at the 2009 Immersion Lithography Symposium, Prague CZ.
7. S. Chang *et al.*, Proc. SPIE **7275**, 72750D (2009).