Estimation of layout densities for CMOS digital circuits

F. Moraes[1], L. Torres[2], M.Robert[2], D.Auvergne[2]

[1]PUC-RS - Av. Ipiranga, 6681 – Prédio 16
90619-900, Porto Alegre, Brazil
e-mail: moraes@andros.inf.pucrs.br

[2]LIRMM, UMR 5506 CNRS/Uni. Montpellier 2
161 rue ADA
34392 – Montpellier – Cedex 5 - France
e.mail : {torres,robert,auvergne}@lirmm.fr

Abstract

The transistor density is one of the parameters to be considered for an optimal use of CMOS process. Therefore, layout strategies have to be evaluated through metrics considering all the involved parameters. The objective of this paper is to study the real transistor density available for a given technology at the cell and circuit levels. This study has three main interests: (i) to evaluate the quality of the layout synthesis tools in terms of area; (ii) to give a feedback to the logic synthesis step, allowing an accurate area prediction from the gate level abstraction; and (iii) to determine a transistor density roadmap.

1. Introduction

The availability of submicronic process with multi-layer routing possibilities opens the way for defining new optimal layout strategies. The routing, shielding and spacing of multi-level signals must conserve signal integrity [1]. All these considerations must be used in defining layout strategies and styles of implementation.

Layouts of random logic circuits can be synthesized using methods like full-custom design, standard-cells, or automatic layout generation. Layouts using full-custom design are extremely dense, however due to the high time to market constraint this approach is practically not used anymore. The standard-cell approach is currently the solution used to synthesize random logic blocks, from a library of full-custom cells automatically placed and routed. The automatic layout synthesis offers some advantages over standard-cells, such as easier technology migration and transistor sizing, paying a higher cost in area.

In order to fill the gap left by the non definition of metrics to evaluate layout densities and to predict the trends of future processes, this paper proposes a method to estimate transistor densities associated to layout strategies. The basis of this method is a regular layout style, associated to assumptions on required routing area and transistor sizes. This approach is used to give an overview on the evaluation of transistor densities for submicronic CMOS technologies. The main references used to compare our results are extracted from the National Technology Roadmap for Semiconductor report [2] and from semiconductor vendors data [3].

The paper is organized as follows. Section 2 presents some characteristics of the layout style used to estimate transistor densities. Section 3 presents the method to

estimate area. Section 4 compares the densities obtained from layout synthesis and standard-cell layout methodologies to that defined in the previous section.

2. Layout Style

The layout style used for density estimation is the ***linear-matrix*** [4] style: all transistors are placed into two parallel diffusion rows. The main cost function is the diffusion gap minimization, aiming at area and side-wall capacitance reduction. A cell with no diffusion gaps has minimal area and minimal parasitic capacitance values. Figure 1 presents a linear-matrix style adapted for submicronic technologies.
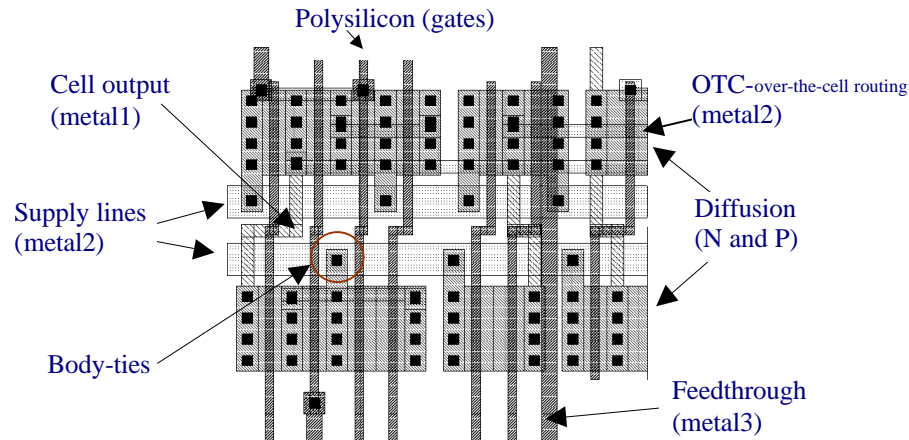

Figure 1 – Linear-matrix style, with three metal layers and stacked contacts

The main features of this style are: over-the-cell routing (placing stacked contacts over drains/sources), supply lines between transistors (with body-ties placed over supply lines), four layers for routing between rows (vertical polysilicon/metal3 and horizontal metal1/metal2), griddles routing, metal3 over cells connecting nets at non adjacent rows. If silicides are used over diffusion, it is possible to minimize the number of contacts in drain/source regions, reducing the constraints imposed to the OTC routing [5].

3. Performance prediction

Our approach consists in determining the reliability of silicon foundries density measures, and to define a method to evaluate there densities from basic design rules.
The target layout style is the multi-row linear-matrix, currently used by layout generators like TROPIC [6] and LAS [7]. Cells are placed into parallel rows, similarly to the cell-based approach. Two main topologies can be used for the multi-row approach (Figure 2): channel-based and abutted rows, depending on the adopted strategy for supply routing, with or without duplication of metal supply layer.
The layout height is determined by the design rules and which defines the minimal distance between 'N' and 'P' transistors and the placement of the body-ties (contacts to the substrate). The layout width is defined by the contacts to the drains and sources

and the minimal diffusion area. As our objective is to evaluate the maximum transistor density, we choose the abutted-row topology, since in this approach the supply lines are not duplicated.

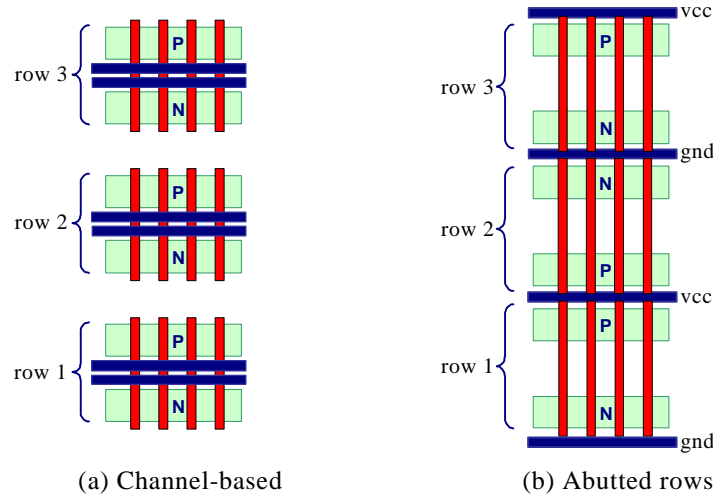(a) Channel-based   (b) Abutted rows

Figure 2 – Multi-row approach

For an N-well process, the layout for a transistor pair (like an inverter) is presented in Figure 3a. In this layout, the connection in the drain region is implemented with contacts and by abutment in the source region. In this way, an even number of P or N transistors can be implemented by mirroring the initial pair, as represented in Figure 3b.

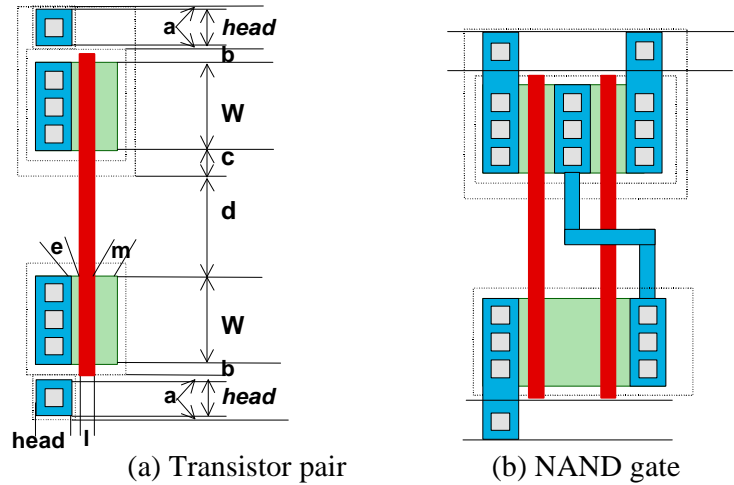(a) Transistor pair   (b) NAND gate

Figure 3 – Linear-matrix layouts

The area of a linear-matrix cell, without routing, is a function of the following design rules:

- **a**: spacing between diffusion and implant, for body-ties;
- **b**: spacing between diffusion and implant, for transistors;
- **c**: spacing between P diffusion and N well;
- **d**: spacing between N diffusion and N well;
- **e**: diffusion contact spacing to polysilicon gate minus overlap of contact;
- **l**: transistor length (defines the technology);
- **m**: maximum of (spacing between two polysilicon gates) and (½.head + e);

- **head**: contact width plus 2 times the overlap of contact;
- **width**: average transistor width in the circuit.

For the abutted-row approach, the transistor density (tr/mm^2), for n rows can be developed as:

- **Transistor number**: $tr\# = 2.n$, (a transistor pair per row);
- **Layout height**:
$$height = n.(c+d+2.w)+(n-1).(head+2.a+2.b)+2.(head+2.a+b)$$
$$height = n.(c+d+2.w+head+2.a+2.b)+head+2.a$$
- **Layout width**: $width = head + e + l + m$

$$density = \frac{Tr\#}{height * width} \tag{1}$$

The equation (1) gives the *optimal* density for the linear-matrix style to properly evaluate transistor densities. Two important points must be analyzed : routing area and transistor size.

The routing area, for design rules with two metal layers, is about 50% of the chip area. For design rules with three metal layers, it is possible to aggressively use the over-the-cell routing [5] to reduce the routing area. There is no consensus in how much area is used in these processes for routing. However, if we consider that it is possible to reduce to half the routing tracks, we can assume the routing area for designs with three metal layers equals to 25% of the chip area. The approach "zero routing footprint" [8] (no routing area) is possible because the cells are designed to be transparent to several routing tracks, increasing in this way the cell height. When more than three metal layers are available for routing, they must be used to connect different blocks, not inside them (transparency over blocks).

The circuit can be optimized for area, for power, for delay, or for some trade-off between these functions [9][10]. As shown in Figure 4, sizing transistors at the minimum width results in smaller area and "power", with delay controlled by loads (diffusion and routing capacitance). For larger width (w>16µm) the delay is almost independent of the load variations and is only dominated by the diffusion capacitance, (which is proportional to the transistor width). Intermediate solution can be obtained by optimizing the delay with locally reasonably sized transistors (area-delay trade-off). Currently, standard-cells are sized for delay optimization, since the load supplied by the cells is unknown. Some libraries have different templates for each cell, one for each optimization cost function. The price is the cost to manage and maintain such libraries. This constraint does not exist for layout synthesis tools: each cell can be sized according to the output load.

We simulated several circuits with transistors sized following different objectives :
- sizing for area at the minimum width (W_{min}),
- sizing for area-delay trade-off at $8W_{min}$,
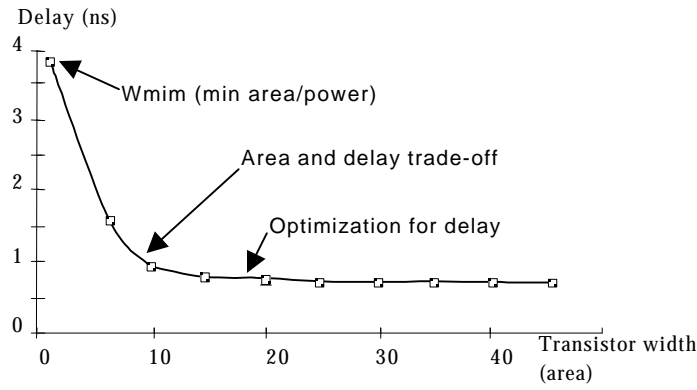- sizing for delay optimization at $16W_{min}$.

Figure 4 – Area and delay trade-off

Figures 5 and 6 show the transistor densities for 3 design rules [11] (units are in µm):

- l=0.8 µm (DML): *a=0.4, b=1.0, c=2.4, d=4.8, e=0.2, l=0.8, m=1.2, head=2.0, W_{min} = 1*
- l=0.6 µm (DML): *a=0.3, b=0.75, c=1.8, d=3.6, e=0.15, l=0.6, m=0.9, head=1.5, W_{min} = 0.75*
- l=0.5 µm (TML): *a=0.3, b=0.6, c=1.5, d=3.0, e=0.1, l=0.5, m=0.7, head=1.2, W_{min} = 0.7*

Figure 5 plots the maximum densities obtained from equation 1; in curves presented in figure 6 we consider the routing area with two and three metal layers available for routing (maximum density multiplied by 0.5 and 0.75 respectively).
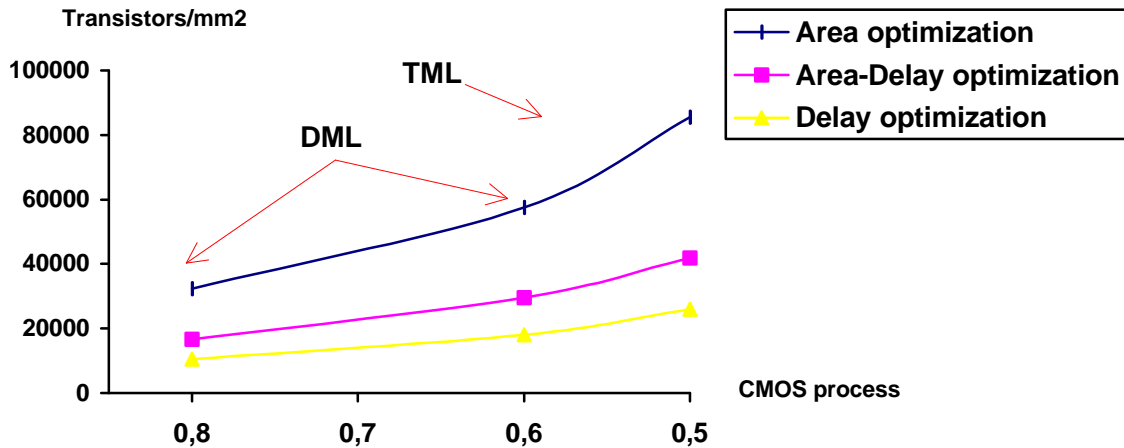

Figure 5 : Densities (tr/mm^2) for different CMOS processes estimated by equation 1

(DML = dual metal layers and TML = three metal layers)

Values presented in Figure 5 and 6 show the transistor densities trends for submicronic technology for DML (Double Metal Layer) and TML (Triple Metal Layer) layout synthesis tools considering different transistor sizing strategies. As we discussed below, these data can be used as the basic elements to the determine transistor density roadmap, to compare with silicon foundries, and to evaluate layout synthesis tools.

The difference between the 0.8 µm and 0.6 µm processes used is a 0.75 shrinking factor 0.75. Consequently, for the same transistor number, the area was reduced by 0.56 (0.75^2) and the density increased by 1.78 (1/0.75^2). Applying a 0.7 shrinking

factor in the 0.5 µm process gives the density of the 0.35µm CMOS process. Prediction for the 0.25 µm and the 0.18 µm can be obtained in the same way. With a 0.7 shrinking factor, the transistor density is multiplied by 2 ($1/0.7^2$).
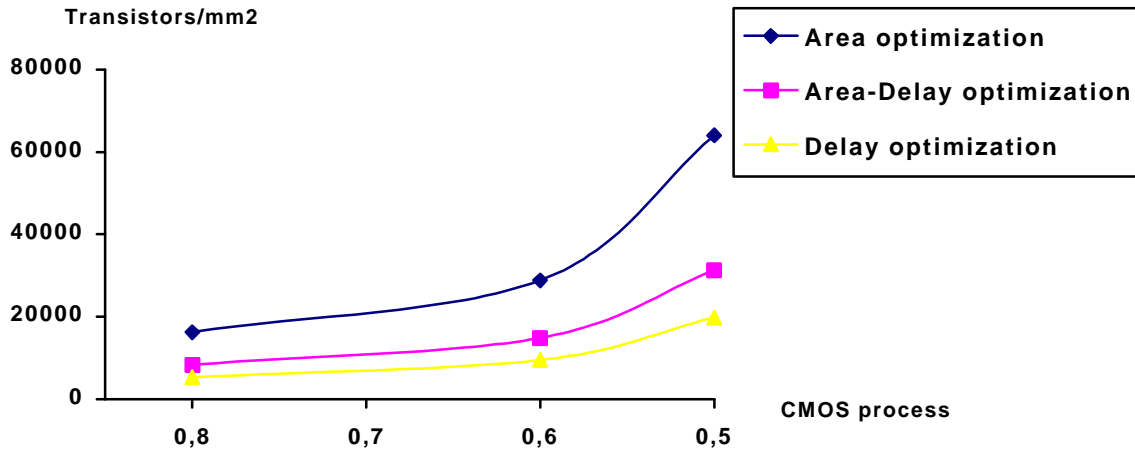

Figure 6 - Densities (tr/mm^2) for different CMOS processes estimated with routing

Using as a reference the 0.5 µm process density with three metal layers for routing and a sizing area-delay optimization, we can extrapolate the densities for different processes (Table 1).

Process (µm)	Transistor Density
l=0.8	8296
l=0.6	14748
l=0.5	31332
l=0.35	62644
l=0.25	125328
l=0.18	250656

Table 1- TROPIC densities Roadmap (tr/mm^2)

These data can be used :
- to evaluate the quality of layout synthesis tools, comparing the layouts obtained from a layout generator to the values obtained from the tables, area used for routing are deduced,
- to predict the area to be used by a circuit, from the number of equivalent gates and the design rules; this information can be very useful for logic synthesis tools to evaluate the area for a given cost function,
- to analyze the trends in terms of transistor area and transistor count for the next generation of integrated circuits.

For example, applying twice a shrinking factor 0.7 over the 0.18 µm process results in transistors with l=0.07 µm, with 1,000,000 transistors per square millimeter. Moreover, the integration of an IC with 10 cm^2, will result in circuit with 1 billion transistors, which is not too far from the estimation of the SIA roadmap [2].

4. Results

4.1. Layout synthesis tools

Table 2 shows the densities obtained from different benchmarks implemented with layout synthesis tools (LAS[7] and TROPIC[12]). The main characteristic of these tools is the absence of pre-characterized cells, with on-the-fly cell generation. The columns "Tr#" and "Rows#" indicate the number of transistors in each circuit and the number of rows used to implement the final layouts respectively. The columns LAS and TROPIC2 give the densities for layouts synthesized with two metal layers for routing. TROPIC3 is the TROPIC version for routing with three metal layers [12]. The circuits were generated using the 0.7 μm process, with transistors sized for the area-delay trade-off.

Circuit	Tr#	Rows#	LAS	TROPIC2	TROPIC3
Adder	28	2	4780	5942	12111
Addergate	40	2	5598	6020	11176
Alu	260	4	5812	5294	9061
Alugate	432	4	6050	5405	8592
Rip	448	5	5961	5610	10207
Cla	528	5	5614	5498	9146
Hdb3	570	4	4903	6516	7758
Mult6	972	7	5950	5779	7994
Mult2	4512	16	4879	4080	7057
Average transistor density:			*5505*	*5571*	*9233*
Estimated transistor density (Figure 6)			*8296*	*8296*	*12444*

Table 2- Transistor densities (tr/mm^2), 0.7 μm process, w=8μm, l=0.8μm

We observe a more important contribution of the routing area than expected comparing the densities obtained from layout synthesis with the predicted ones. As mentioned before, the routing area should be respectively 50% and 25% for 2 and 3 metal layers. We got 66% and 44%. This difference can be explained by several reasons, at the layout level:
- over-the-cell routing is not used in the two metal process;
- stacked contacts are not allowed (two-metals process);
- the great number of contacts in each drain/source prevents routing over the transistors;
- large supply lines increase the circuit height;
- as the area of the block corresponds to its bounding box, rows with different lengths produce empty spaces in the layout;
- the cells are separated by diffusion gaps, increasing the circuit width;
- the length of the cells will increase when diffusion gaps are necessary;

- the compactor used to translate the symbolic description into layout also plays an important role, for example, if jogs are automatically inserted the final area as expected to reduced.

Thus, the layout synthesis tools must improve the efficiency of the layout style, (placement and routing), in order to achieve the estimated densities, when routing is considered.

4.2. Standard-cell approach

At the cell level, the densities directly evaluated on standard-cell layouts of an industrial library (CMOS processes 0.7 µm and 0.5 µm.) are 12500 and 54000 tr/mm^2 respectively. These densities values are not different from the estimated solution for area without routing (16197 and 64171 tr/mm^2 (Figures 5 and 6). The cells of a library are handcrafted (full-custom design), with no pre-defined position for transistor placement and no preferential direction for the routing layers. At the circuit level, the effect of routing is quite minimized.

Circuit	Tr#	Gates#	Density
Register	242	60	6454
Measure	3545	861	7082
Valid	5993	1498	7758
Top	6417	16	7627

Average transistor density: *7230*

Table 3 - Density (tr/mm^2) for circuits generated from standard-cells (0.7 µm process)

The advantage of standard-cells over automatic layout generation consists in the possibility in optimizing complex functions at the cell level, such as flip-flops and adders. For example, the circuit "valid" has 35% of the transistors used to implement flip-flops. When flip-flops are generated using the linear matrix style, all transistors must be placed horizontally and the internal connections must respect the preferential directions for routing. On the other hand, when flip-flops are designed by hand, transistors can be freely placed, with no restrictions for the internal routing.

To automate the creation of cell layouts competitively with handcrafted quality layouts is a non-trivial problem. To date, the real complexity of the cell synthesis process has been the main obstacle to its commercial success [13]. Besides the problem of spending more area than standard-cells, the automatic layout synthesis is a very attractive solution to synthesize random logic blocks. Main reasons can be advanced for that :

- the number of cells available into a library is quite small when compared to the possibilities given in using complex gates, different functions can be implemented with complex gates with less than four serially connected transistors is 3503 [14],

as shown in [15], it is possible to reduce the overall number of transistors up to 40%, reducing in this way area, delay and power,

- technology independent synthesis is necessary in supporting rapidly advancing processes [16] and layout generation permits an easy technology migration, allowing to quickly migrate from a set of design rules to another one,
- transistors can be individually sized according to the designer constraints, with the possibility to adjust the delay of the critical path(s), downsizing the transistors outside the critical path, in order to achieve power-delay optimization,
- the cost for cell and library maintenance is very expensive, with a *virtual library* concept (no pre-characterized cells) this cost is suppressed.

4.3. Roadmap

Figure 7 compares different transistor densities roadmaps : silicon vendors [3], SIA [2] (Semiconductor Industry Association) for ASIC and microprocessors, estimated maximum density (equation 1), estimated density with Tropic generator and some density results coming from industrial standard cells libraries.
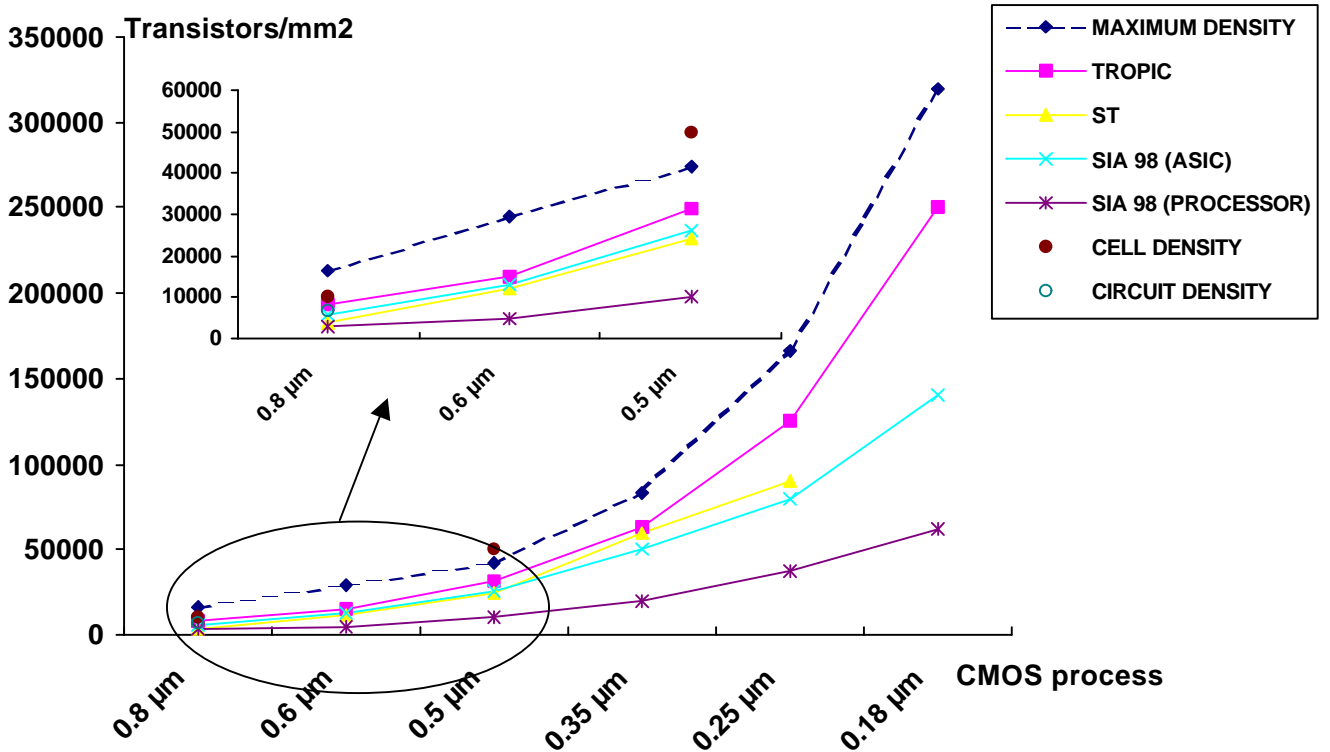

FIGURE 7 : Densities roadmap comparison

Respectively, the SIA roadmap and the maximum density (equation 1) appears as a lower and upper bound. The SIA roadmap is quite conservative, since it is obtained from the number of functionally active transistors/mm^2 (with no pad area consideration) designed by automated layout tools [2]. The SIA processor roadmap is a particular curve showing the embedded memory effect on chip density for complex processors (11 millions of transistors per chip for a 0.25μm process).

Between these limits, transistor densities of layouts corresponding to random logic generated by TROPIC and some examples obtained from silicon foundries (SGS-Thomson, cell and circuit density) are presented. The cell density value appears not too different of the upper bound, however, at the circuit level the routing between gates reduces the density to values closed to the SIA values.

5. Conclusion

The contribution of this paper is the definition of a method to estimate silicon area to implement layouts (regular and standard-cells layouts), starting from the number of transistors and the layout design rules. Routing area was considered as a constant percentage of layout area, 50% and 25% for two and three metal layers respectively. Three sizing possibilities were considered: area, area-delay and delay optimization. The data obtained from this method can be used to evaluate layout synthesis tools, to estimate area in logic synthesis tools and to predict the densities of new processes.
The difficulty for CAD developers is to know how far the heuristic algorithms for cell synthesis, placement and routing are from the optimal. The data presented in equation (1) can be used as an upper bound for transistor densities, driving heuristic algorithms to give results near to predicted values. Considering the number of routing layers available in actual processes (6 or 7), great effort must be given in defining strategy for local and inter-block connection and signal propagation.

6. References

[1]   H.T. Heineken, W. Maly. " Performance-Manufacturability Tradeoffs in IC Design", DATE'98, Paris 1998.
[2]   Web site : http://www.sematech.org.
[3]   Web site : http://www.st.com.
[4]   T.Uehara, W.van Cleemput. "*Optimal Layout of CMOS Functional Arrays*". IEEE Transactions on Computers, Vol. C-30, No. 5, May 1981, pp. 305-312.
[5]   J.Kim; S.M.Kang. "*A New Triple-Layer OTC Channel Router*". IEEE Transactions on CAD, Vol. 15, no. 9, September 1996, pp. 1059-1070.
[6]   F.Moraes, N.Azemard, M.Robert, D. Auvergne. "*Flexible macrocell layout generator*". 4th ACM/SIGDA Physical Design Workshop, Los Angeles, 1993, pp. 105-116.
[7]   CADENCE. "*Virtuoso Layout Synthesizer - LAS - User Guide*". CADENCE™ Version 4.2, October 1991.
[8]   N.Sherwani. " *Algorithm for VLSI physical design automation*". Kluwer Academic Publisher, 538p. 1995.
[9]   S.S.Sapatnekar; S.Kang. "*Design Automation for Timing-Driven Layout Synthesis*". Kluwer Academic Publisher, 296p. 1993.
[10]   S.Turgis, N.Azemard, D.Auvergne. "*Design and selection on buffers for minimum power-delay product*". ED&TC'96: The European Design & Test Conference, Paris, France, March 11-14, 1996, pp. 224-228.
[11]   ATMEL-ES2. ECAT05 and ECPD07 design-rules, 1998.
[12]   F.Moraes; R.Reis; F.Lima. "*An Efficient Layout Style for Three-Metal CMOS Macro-Cells*". VLSI'97, Gramado (Brazil), pp. 415-426.
[13]   M.Lefebvre; D.Marple; C.Sechen. "*The Future of Custom Cell Generation in Physical Synthesis*". DAC'97.
[14]   E.Detjens; G.Gannot; R.Rudell; S.Vincentelli; A.Wang. "*Technology mapping in MIS*". ICCAD, Santa Clara, 1987, pp. 116-119.
[15]   A.Reis; M.Robert; D.Auvergne; R.Reis. "*Associating CMOS Transistors with BDD Arcs for Technology Mapping*". Electronic Letters, Vol. 31, No 14, July 1995.
[16]   M.Guruswamy; R.L.Maziasz; D.Dulitz; S.Raman; V.Chiluvuri; A.Fernandez; L.G.Jones. "*CELLERITY: A Fully Automatic Layout Synthesis System for Standard Cell Libraries*". DAC'97.