Density and Reliability Predictions for a General Logic Structure for Custom LSI

M. Sievers
Communications Systems Research Section

A general logic structure (GLS) for implementing arbitrary functions in integrated circuits has been described in a previous report. Density and reliability predictions for the GLS will be presented in this article. The GLS has been found to be more dense than programmed logic arrays (PLA) and certain configurations of "optimized" macros. Macro is used here to mean a predefined function that may be inserted into a design.

A reliability model is presented that includes the possibility of undetected manufacturing flaws. This model is more accurate than models that consider only so-called wear-out failures. It may be used to indicate how much preinstallation test coverage is necessary to guarantee a given installed reliability.

I. Introduction

A general logic structure (GLS) for constructing integrated circuits has been reported on in Ref. 1. The GLS is a two-dimensional array of wires into which logic functions are mask programmed. A stick diagram of an unprogrammed NMOS GLS is shown in Fig. 1. The GLS consists of metal columns (dot-dashed lines), grounded diffusion rows (dashed lines), and polysilicon implicant rows (solid lines). Metal column density in Fig. 1 is twice that reported in Ref. 1. Alternate columns are used to build gates, the other columns are used to carry signals or power.

Figure 2 shows the stick diagram of a programmed NMOS GLS. Enhancement mode transistors are created at the intersection of polysilicon and a diffused region. The polysilicon becomes the transistor gate and the diffusion forms the contacts to the channel. A NOR logic gate is created by connecting one or more transistor channels between a gate column and ground. When any of the transistors conduct, the gate output is zero. Pull-up resistors pull a gate column to the positive supply voltage when no pull-down transistors are conducting. The pull-up transistors are made from depletion mode transistors. Depletion mode transistors are formed by ion implanting the transistor channel prior to the polysilicon and diffusion steps. The GLS is programmed by creation of pull-up and pull-down transistors, cuts, and contact points.

The philosophy behind the GLS structure is to optimize wiring regularity while sacrificing logic regularity if necessary. This approach is different from the usual in which logic is optimized at the expense of wiring. It turns out that wiring area consumes the larger portion of chip area because logic structures can easily be made compact. It can be expected, therefore, that a structure that optimizes wiring area will often

occupy less area than a structure that optimizes logic area. Density estimates for the programmed logic array (PLA), GLS, and various configurations of "optimized" macros are given in Section II.

Section III will deal with reliability estimates for the GLS. A model is presented that includes the possibility of undetected circuit flaws occurring in an installed chip.

It is common to model integrated circuit failure only in terms of the arrival of random wear-out failures such as metal separation due to electromigration. A Poisson arrival is assumed for the failures. Using this model, the probability of k failures occurring in a time interval t is

$$P_k = (\lambda t)^k e^{-\lambda t}/k! \quad (1)$$

where λ is the failure rate. From Eq. (1), the probability that no failures occur in the interval is

$$P_0 = e^{-\lambda t} \quad (2)$$

This model is accurate when chips are 100-percent tested prior to installation. This is not always possible and may not be necessary as long as the consequences of reduced testing can be determined. The model in Section III predicts the effects of less than perfect testing.

II. Density Estimates

A good measure of the efficiency of a chip design is its gate density. In this section, the gate density for three design methods are estimated and compared.

An estimation of the gate area required to place and wire a given number of gates on a chip can be made by slightly recasting the question posed by the wireability analysis in Ref. 2. That article considers the probability of successfully wiring a given number of gates in a given area. The question considered here is: what area is necessary to guarantee the wireability of a given number of gates? The analysis in Ref. 2 is pessimistic and the area estimates computed here based on that paper will be pessimistic.

Before beginning the analysis, a few definitions are required. Define pitch as the separation between gates. The average pitch is denoted by $\bar{R}$. Demand is defined as the length of wire, measured in pitches, required by a given circuit. Finally, define capacity as the total wire length available in a circuit.

Demand, D, is computed by the equation

$$D = \text{No. of gates} * \bar{R} * \text{fan-in/gate} \quad (3)$$

Wireability can be assured if the capacity is twice the demand.

Let the number of gates to be placed on a chip be m. Assume these are arranged in a square pattern with $\sqrt{m}$ gates on a side. Consider the GLS in which f is the fraction of rows and gate columns available for wiring. Vertically there will be

$$V_w = \sqrt{m} + \sqrt{m}f + w_v \quad (4)$$

wires per row where w_v is the number of wires needed vertically to guarantee wireability. If there are r pitches per column, the total number of vertical pitches is

$$V_p = r(m(1+f) + w_v)/2 \quad (5)$$

since there are two vertical wires per pitch. Similarly, the number of horizontal pitches is

$$H_p = c(2\sqrt{m}f + w_h)/4 \quad (6)$$

where w_h is the number of horizontal wires needed for wireability and there are c vertical pitches per row. There are four horizontal wires per horizontal pitch, so the number of rows r is

$$r = \sqrt{m} + w_h/4 \quad (7)$$

and the number of columns c is

$$c = \sqrt{m} + w_v/2 \quad (8)$$

Summing vertical and horizontal capacity, letting $w_v = w_h = w$, and doing some algebra, the number of wires needed in each direction is

$$w = ((8D - 2m(1+f) + (\sqrt{m}/4(7+3f))^2)^{1/2} - \sqrt{m}/4(7+3f) \quad (9)$$

For an NMOS GLS based on the Caltech design rules (Ref. 3), the horizontal pitch is $14\lambda_d$ and the vertical pitch is $36\lambda_d$

where is λ_d the minimum defineable distance. Total area is therefore

$$A = (14(\sqrt{n} + w/2)\,(36(\sqrt{m} + w/4)) \qquad (10)$$

Table 1 tabulates the demand, w, and GLS area for several numbers of gates. The value for $\bar{R}$ was taken from Ref. 2, fan-in is assumed to be 2.5, and f was set to 1/4.

To compute the area requirements for a programmed logic array (PLA) of m gates, consider Fig. 3. The area labeled PU is the pull-up resistor area. Half of the m gates are located in the upper AND plane, and the remaining half are located in the bottom OR plane.

Using the Caltech NMOS design rules, PLA implicant rows are an average of $7\lambda_d$ apart, as are the columns. A depletion mode transistor is $18\lambda_d$ long. All of the wires needed by the PLA are internal to it in the form of large fan-in AND and OR gates. Let the number of input wires be w where

$$w = n(m/2) \qquad (11)$$

and $n > 0$. From Fig. 3, the area is

$$A_{\text{PLA}} = (n(m/2)\,(7) + 18)\,(m/2{*}7) + (m/2{*}7)\,(m/2{*}7 + 18) \qquad (12)$$

Simplifying

$$A_{\text{PLA}} = 12.25nm^2 + 12.25m^2 + 126m \qquad (13)$$

Table 2 compiles PLA areas for several values of n and m.

The final area estimates are based on the structure shown in Fig. 4 and shows a collection of optimized modules wired together in a grid pattern. These modules are very dense and do not permit wires to run through them. The number of input-output terminals, T, in each cluster can be predicted via Rent's rule (Ref. 2):

$$T = Am^{\text{p}} \qquad (14)$$

where A is the fan-in of the internal gates, m is the number of internal gates, and p is a fraction, $1/2 \leqslant p < 1$, that is related to the relative function per pin. For $p = 1/2$, the number of pins per function is small and most of the wiring is internal to the module. As p approaches 1, the number of terminals becomes the sum of the internal gate fan-in and most of the wiring is external to the module. Assume a "tight" design where $p = 1/2$ and there are 2.5 inputs per gate. The required number of terminals is from Eq. (14):

$$T = 2.5\,m^{1/2} \qquad (15)$$

which is tabulated in Table 3.

Let the area occupied by m gates and connecting wires in an optimized module be half of what would be required by the equivalent number of GLS gates excluding wires. Also, assume the optimized modules are square. Table 3 lists the optimized module area for several values of m.

Tables 4 through 9 list demand, and area for various configurations of optimized modules, and gates per module. Area computations are made following the procedure outlined for the GLS.

Area estimates tabulated in Tables 1, 2, 4, and 7 are plotted in Fig. 5. From Fig. 5, it can be seen that the GLS does better than the PLA or clusters of optimized design for a sufficiently large number of gates. Another conclusion that can be made is that modules should be designed so that some number of wires can pass through them. Reducing the number of blocked wire paths reduces the overall chip area.

III. Reliability Estimation

Reliability estimates based on Eq. (1) assume that chips are 100 percent tested prior to installation. This level of testability is usually very costly and is becoming decreasingly possible as the number of devices per chip increases. A reliability model should include the effects of less than perfect chip testing.

Assume that clustered flaws are of such serious nature that they are always 100 percent detectable. The remaining flaws are thus randomly distributed across the surface of the wafer. If the flaw density is N flaws per unit area, then the probability of no flaws in an area A is

$$P_0 = \text{e}^{-NA} \qquad (16)$$

The probability of at least one flaw is

$$P(k \geqslant 1) = 1 - \text{e}^{-NA} \qquad (17)$$

Let the ratio of active circuit area to total area be denoted r. Then the probability of a random circuit flaw is

$$P_{cf} = r(1 - e^{-NA}) \tag{18}$$

Denote by $\overline{TC}$ the probability that a flaw goes undetected where $\overline{TC}$ is one minus the test coverage. Then the probability of an undetected circuit flaw is

$$P_{ucf} = \overline{TC} P_{cf} \tag{19}$$

and the probability that no such flaw exists is

$$\overline{P_{ucf}} = 1 - P_{ucf} \tag{20}$$

The occurrence of a random undetected circuit flaw may be considered independent of the occurrence of wear-out failures. Therefore, reliability may be computed by the product of the probability of no wear out and the probability of no undetected circuit flaws.

$$R = e^{-\lambda} \cdot \overline{P_{ucf}} \tag{21}$$

Equation (21) may be used to balance chip area, initial testing, r, and various redundancy schemes to achieve the desired reliability.

The reliability of a triple modular redundant (TMR) system built on a single GLS chip is plotted in Fig. 6 for several variations of area and testing. The voter and associated wires are assumed to occupy the area of one module. The ratio r is 0.7 for the GLS. TMR reliability is computed by

$$R_{TMR} = R_v R^3 + \binom{3}{2} R_v R^2 (1 - R) \tag{22}$$

where R_v is the voter plus wire reliability and R is computed by Eq. (21). Voter reliability is assumed equal to module reliability since these occupy the same area.

Testing of a very large chip will probably be done via self-testing techniques. Self-testing need not have 100 percent coverage to be of value. Equation (21) may be used to predict chip reliability for self-testing chips. Totally self-checking schemes such as (Ref. 4) occupy a considerable area. The result is that manufacturing flaws are more likely. A partially self-checking scheme may therefore outperform a totally self-checking scheme.

References

1. Sievers, M. W., "A General Logic Structure for Custom LSI," *DSN Progress Report 42-50*, Jet Propulsion Laboratory, Pasadena, Calif., January and February 1979, pp. 97-105.
2. Heller, W. R., W. F. Mikhail, and W. E. Donath, "Prediction of Wiring Space Requirements for LSI," *Proc. 14th Design Automation Conf. IEEE/ACM*, New Orleans, June 1977.
3. Mead, C., and L. Conway, *Introduction to VLSI Systems*, 1978, ch. 1, 2, text in preparation.
4. Carter, W. C., A. B. Wadia, and D. C. Jessup, "Implementation of Checkable Acyclic Automata by Morphic Boolean Functions," *Symp. Comp. and Automata*, Polytechnic Inst. of Brooklyn, April 1971, pp. 465-482.

Table 1. Characteristics for several numbers of gates

m	$\bar{R}$	D	W	Area, λ_d^2
36	1.387	250	34	168,084
64	1.590	509	49	331,695
100	1.771	886	66	574,308
225	2.117	2382	111	1,518,993
400	2.410	4820	159	2,996,343
900	2.889	13,001	266	7,927,668
1600	3.276	26,208	382	15,775,452

Table 2. PLA area requirements in λ_d^2

m	Area (n = 0.5)	Area (n = 1)	Area (n = 2.5)
36	28,300	36,288	60,102
64	83,328	108,416	183,680
100	196,350	257,600	441,350
225	958,584	1,268,663	2,173,382
400	2,990,400	3,970,400	6,910,400
900	14,997,150	19,958,400	34,842,150

Table 3. Terminal and area requirements of optimized modules

m	T	Area λ_d^2
6	7	1,512
12	9	3,024
18	11	4,536
24	13	6,048
30	14	7,560
36	15	9,072
42	17	10,584
48	18	12,096
60	20	15,120

Table 4. Gates/module = 6

No. modules	Total No. gates	Demand	Area, λ_d^2
4	24	34	60,348
6	36	51	119,286
8	48	68	197,400
12	72	105	422,082
16	96	140	713,136
36	216	350	3,813,756
64	384	713	14,522,760
100	600	1,240	41,890,500

Table 5. Gates/module = 12

No. modules	Total No. gates	Demand	Area, λ_d^2
4	48	44	106,964
6	72	65	205,350
8	96	87	339,836
12	144	135	726,150
16	192	188	1,310,256
36	432	450	6,453,900
64	768	916	24,361,344

Table 6. Gates/module = 18

No. modules	Total No. gates	Demand	Area, λ_d^2
4	72	53	156,392
6	108	80	309,174
8	144	106	504,432
12	216	165	1,083,138
16	288	220	1,823,024
36	648	550	9,632,604
64	1,152	1,120	36,396,096

Table 7. Gates/module = 24

No. modules	Total No. gates	Demand	Area, λ_d^2
4	96	63	217,872
6	144	94	424,536
8	192	130	738,192
12	288	260	1,502,748
16	384	650	2,530,944
36	864	1,323	13,001,024

Table 8. Gates/module = 30

No. modules	Total No. gates	Demand	Area, λ_d^2
4	120	68	259,572
6	180	101	501,126
8	240	135	826,812
12	360	210	1,770,408
16	480	280	2,977,104
36	1,080	700	15,686,484

Table 9. Gates/module = 60

No. modules	Total No. gates	Demand	Area, λ_d^2
4	240	96	517,860
6	360	144	1,013,526
8	480	192	1,667,016
12	720	300	3,596,148
16	960	400	6,050,064

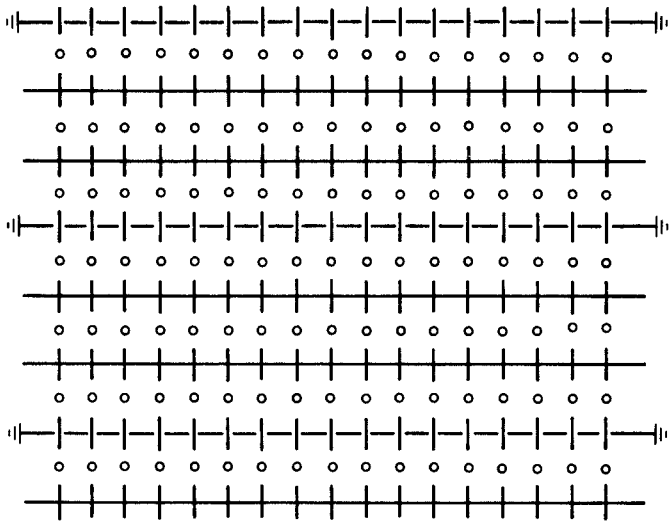

Fig. 1. Unprogrammed NMOS GLS

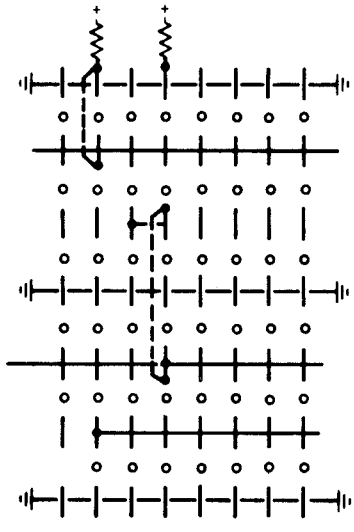

Fig. 2. Programmed NMOS GLS

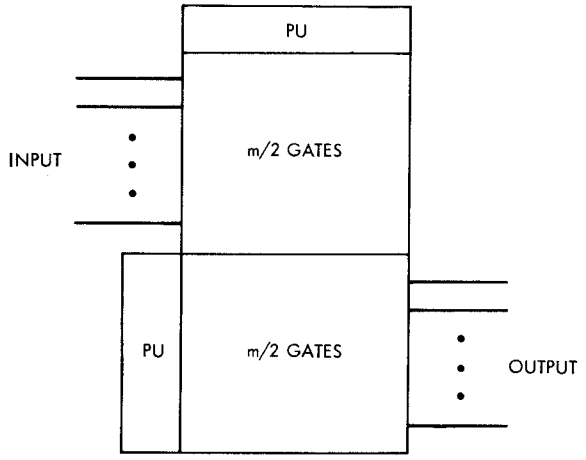

Fig. 3. PLA architecture

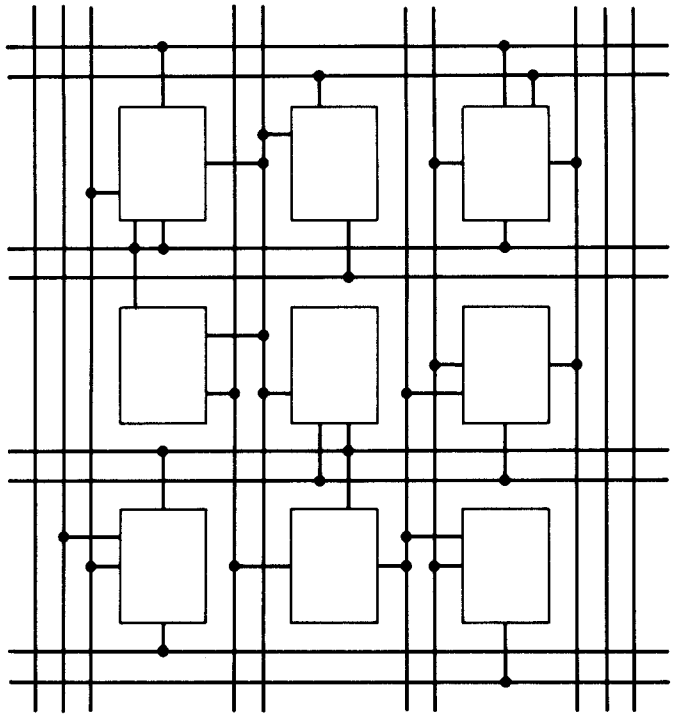

Fig. 4. Optimized macros

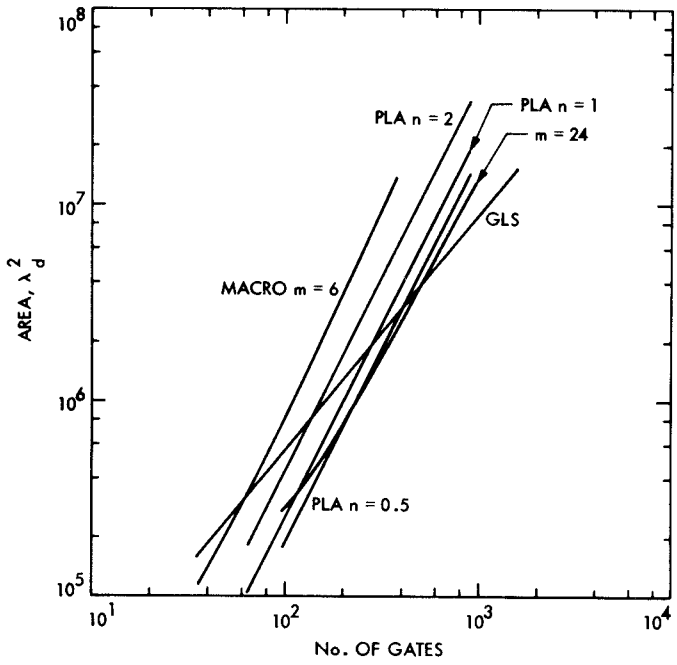

Fig. 5. Area estimates: GLS, PLA, optimized macros

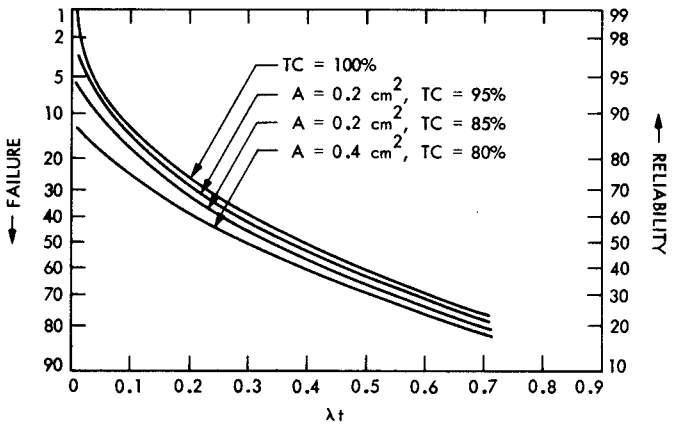

Fig. 6. TMR behavior (N = 2.5 flaws/cm^2)