ELEKTRİK, VOL.1, NO.2, June 1993, ©TÜBİTAK EMO

A Gate Array Architecture For Pipeline Structures

İsmail Enis Ungan and Murat Aşkar
Electrical and Electronics Engineering Department
Middle East Technical University
06531, Ankara, Turkey

Abstract

A new gate array architecture for a pipeline circuit structure is described. The core cells in the gate array are composed of elementary logic gates (AND/NAND and OR/NOR) with which any Boolean function can be implemented. The core cells are designed in the true single phase clock (TSPC) circuit technique and each core cell logic is implemented in a single logic depth. The cells are clocked by a true single phase clock in order to reduce the clock skew. The new gate array architecture is found to be highly applicable to the pipeline and systolic array structures of concurrent systems. A Boolean function and a serial full adder are implemented on the new gate array architecture. SPICE simulations show that 200 MHz clock frequency is reached in 3.0 μm n-well CMOS technology.

1. Introduction

Although the gate arrays cannot fully utilize the use of silicon area, they have always been advantageous over full custom designs when short turn-around time, small design effort and low cost IC production are concerned. In the first generation gate arrays, the cells are separated by field oxide and routing is done over routing channels. The architecture of second generation gate arrays is based on gate isolation technique, in which the cells are separated by transistors whose gates are connected to power lines and routing is done over unused transistors [1]. The gate array IC designs, compared to full custom IC designs, cannot operate at very high speeds. This is because of the high layout parasitic capacitances and resistances.

In order to design high speed IC's layout parasitics should be at minimum. This can only be achieved in full custom designs in which designer has access to all layout masks. In semi custom gate array designs, designer has access only to metalization layers which highly limits the layout design for parasitics minimization. However, a gate array architecture with reduced layout parasitics would show a speed performance close to that of full custom. A gate array structure involving some logic elements as the basic devices, contrary to the classical sense gate array in which the basic device is the transistor, becomes an IC with very fast turn-around time and very high speed. The basic logic elements in this structure should be designed with minimum layout parasitics.

For synchronous system implementations on IC's, the clock distribution has been the major limitation to high speed operation. Clock skew which is known as the overlapping of active clock phases due to delays on the clock signal paths gives an upper bound to the clock frequency and hence to the speed of the system. The clock skew problem gets more complicated when more than single phase clocks are used in the system [2],[3]. In gate array designs, clock skew becomes a hard problem to be solved since the designer cannot easily control the delays on the clock paths.

The true single phase clock (TSPC) dynamic circuit technique in CMOS as suggested in [4], has several superior properties with respect to the other techniques like clocked CMOS [3], domino logic [5], and NORA dynamic CMOS technique [6]. The main property in TSPC is the non-inverted single phase clocking in which no clock skew occurs except for clock distribution delays. Consequently, the system clock rates can be increased and the clock line routing is simplified. The TSPC technique places logic circuits into single phase clocked latch blocks and cascades these blocks to form a pipeline structure. TSPC dynamic logic and TSPC precharge logic are the two kinds of circuit techniques that can be used in a design with single phase clock. Possible TSPC logic designs are shown in Figure 1 [7]. For these reasons TSPC technique is highly suitable for concurrent systems. The concurrent systems implemented by array of logic blocks in a pipeline structure avoid idle times of the logic blocks and therefore improve the overall system performance with very high throughput of data.

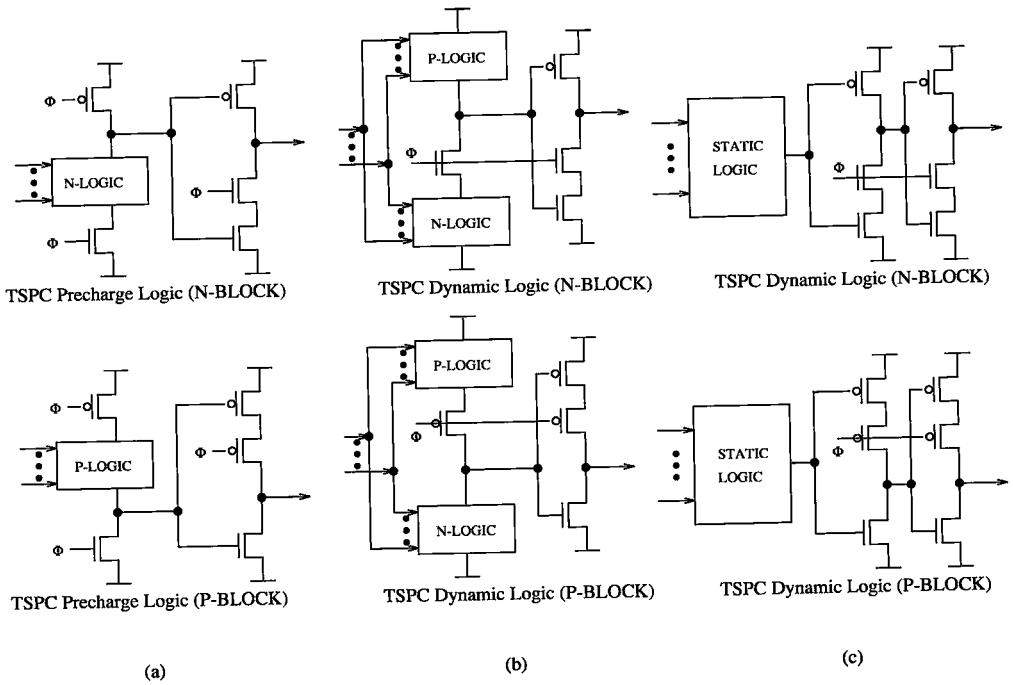
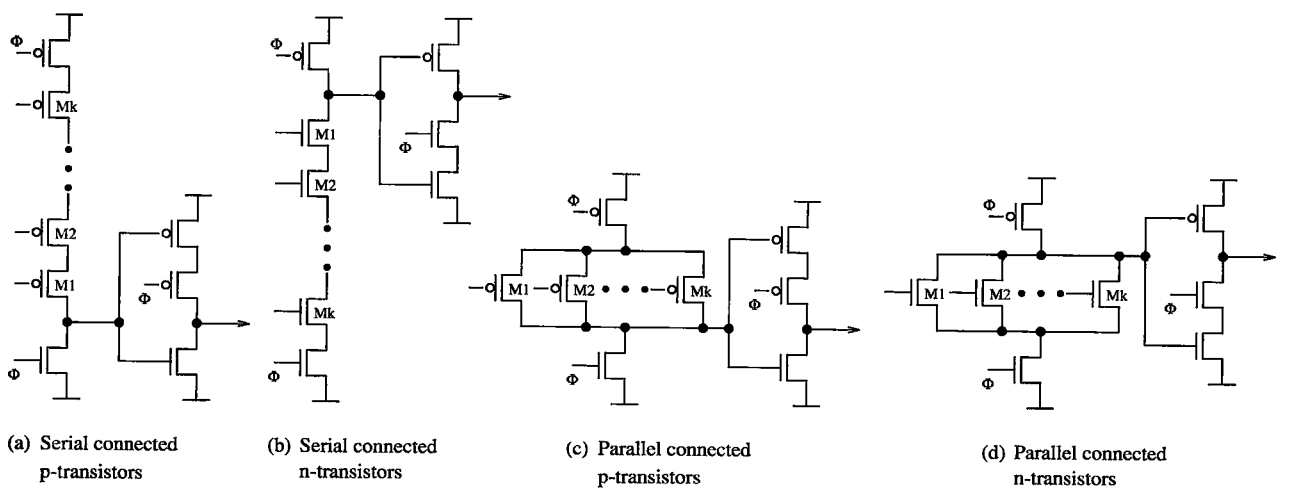


Figure 1. TSPC precharge logic and TSPC dynamic logic examples.

In Section II, we propose a logic style with logic depth one in TSPC technique in order to implement any Boolean function at very high clock frequency with minimum design effort using the new gate array architecture. Core cells and architecture of the gate array are presented in Section III and Section IV, respectively. A Boolean function and a serial full adder on the pipeline structured gate array are given as a design example in Section V. Finally, our conclusions are presented in Section VI.

2. A Logic Style in TSPC Circuit Technique

In a pipeline system design, the system throughput is independent of the number of cascaded pipeline stages and it can be increased by minimizing the delay of each logic block in the pipeline stage. The number of transistors connected in series in a logic circuit degrades the logic delay more severely than the transistors connected in parallel. This is illustrated with a number of SPICE simulations on parallel and serial connected k number of p-transistors and n-transistors in a TSPC precharge logic block. The same number of p-transistors and n-transistors are used in the circuits. The circuits are extracted from the layout so that parasitic capacitances are included in the simulations. The circuit schematics used in the simulations are given in Figure 2. The worst case delay of each circuit is measured between 50 % level of clock and 50 % level of circuit output. The results are normalized with respect to the single n-transistor and are tabulated in Figure 3. It can be concluded from this table that parallel connected transistors have less delay time than serial connected transistors. As the number of transistors increases, the rate of increase in delay time of series connected transistors is larger than that of parallel connected transistors. These are illustrated as plots in Figure 3.



(a) Serial connected p-transistors (b) Serial connected n-transistors (c) Parallel connected p-transistors (d) Parallel connected n-transistors

Figure 2. The serial and parallel connected transistors in TSPC precharge logic technique.

Complex logic function design with the TSPC circuit technique requires extensive amount of design time due to optimization of function implementation (logic style selection), transistor size optimization (device sizing) and layout complexity [7]. In a concurrent system, optimization of the transistor sizes for minimum propagation delay is done for each pipeline stage. As the number of transistors increases, the parasitic capacitances that are dependent on transistor sizes also increase. When a simple circuit is optimized, the increase in the operation speed is not so high. However, when a large and complex circuit is optimized with respect to many parameters, it becomes much faster than its non-optimized form. The logic style with parallel transistors not only simplifies device sizing and layout problems, but also it increases the clock frequency and the layout becomes more regular.

TABULATED NORMALIZED* DELAY TIMES

K	n-transistors		p-transistors	
	parallel	serial	parallel	serial
1	1.00	1.00	3.04	3.04
2	1.07	1.53	3.39	4.71
4	1.22	2.67	4.08	8.40
6	1.36	3.95	4.78	12.57
8	1.50	5.35	5.48	17.29

*Normalized w.r.t. the delay time for n-transistor with K=1.

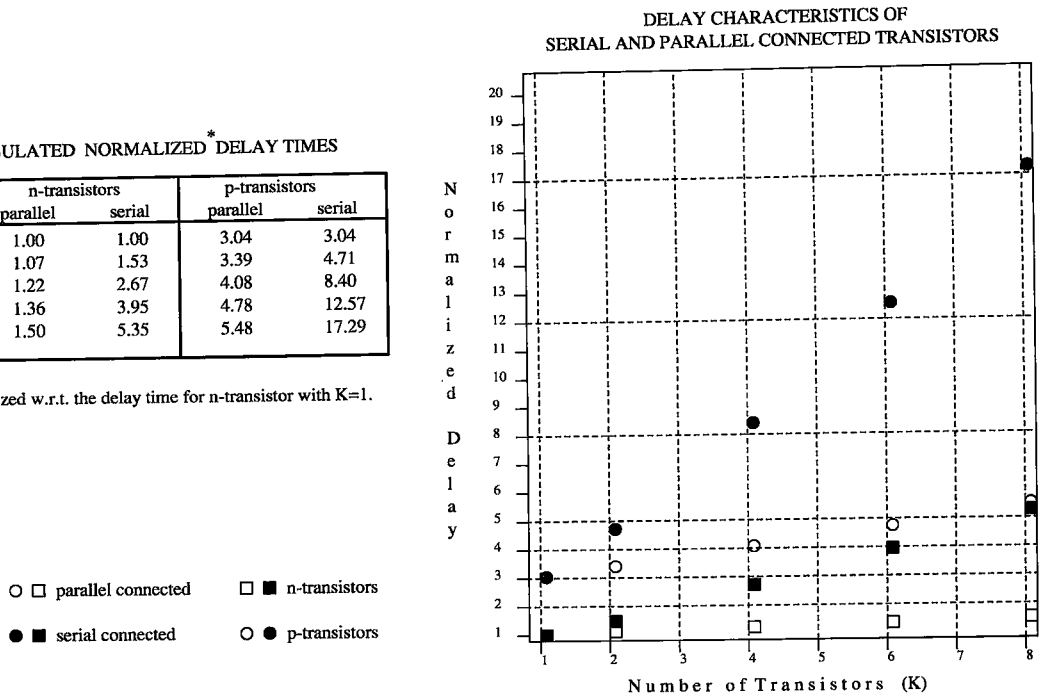


Figure 3. Normalized delay curves for serial and parallel connected transistors.

The logic style in TSPC circuit technique consists of parallel connected k number of p-transistors as the P-logic to implement a k-input AND logic gate , and parallel connected k number of n-transistors as the N-logic to implement a k-input OR logic gate. When the P-logic and the N-logic are embedded in the TSPC precharge logic, they form P-block and N-block, Figure 1. The circuit schematics for k-input AND gate and k-input OR gate are in Figure 2 (c) and 2(d), respectively. P-blocks and N-blocks are arranged one after the other in a pipeline structure. P-blocks precharge while clock is high and evaluation is performed while the clock is at low state, whereas, N-blocks evaluate while the clock is high and precharge while the clock is low. Therefore, the blocks of one type evaluate while the blocks of the other type precharge in the pipeline structure. This is illustrated in Figure 4. In order to implement any Boolean function, an inverting gate is added to the P-blocks (AND gates) and to the N-blocks (OR gates) in the pipeline structure. Since each block has a logic depth of one, due to parallel connected transistors, the clock frequency, hence the function throughput can reach very high values.

As an example, the following 10-input Boolean function F expressed as the sum of minterms is considered:

$$F(A,B,C,D,E,F,G,H,I,J) = \sum(51, 60, 195, 204, 771, 780)$$

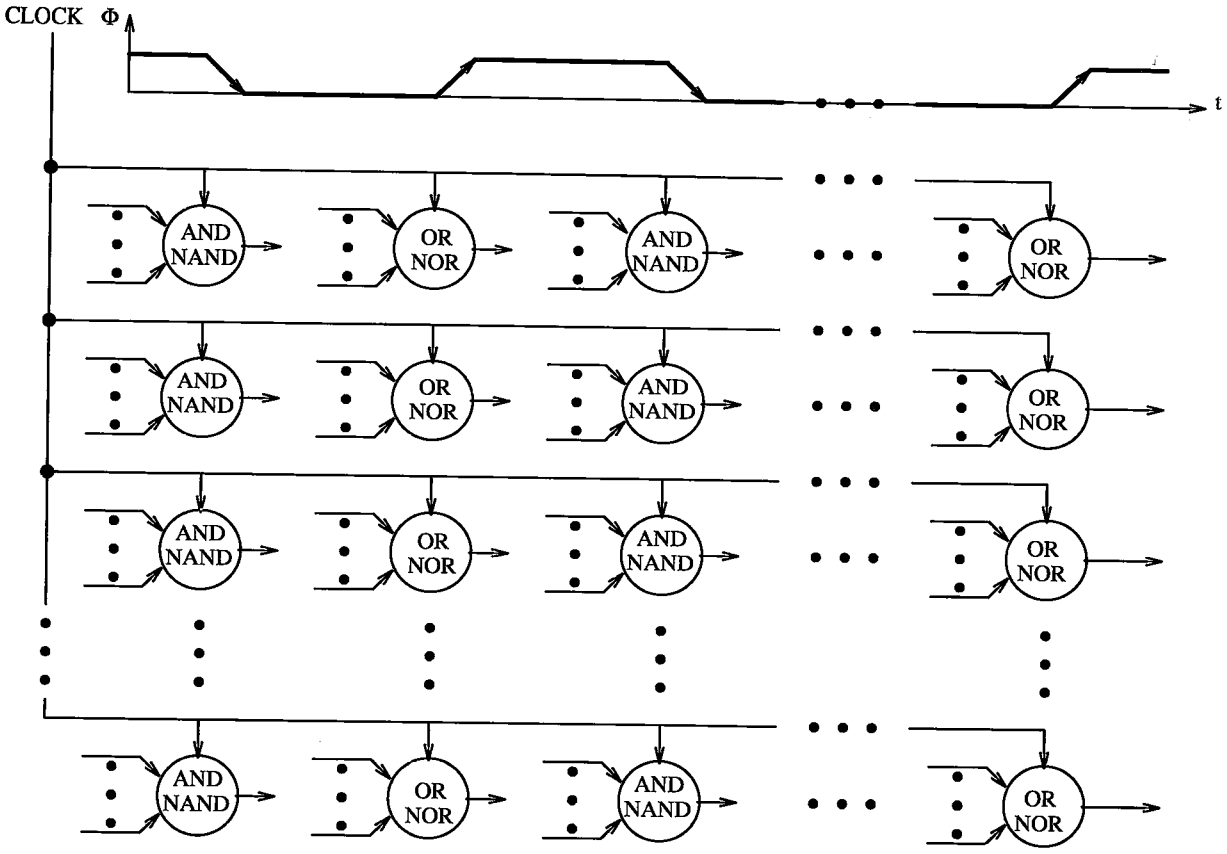


Figure 4. AND/NAND and OR/NOR gates placed on the pipeline structure.

One of the static versions of this function implementation can be by six static AND gates followed by an OR gate in two logic stages as shown in Figure 5(a). The other static version can be by static AND and OR gates in three logic stages as illustrated in Figure 5(b). The corresponding TSPC technique implementations are given in Figures 5(c) and 5(d), respectively. The clock lines are not shown in the figure, and the symbols $\wedge$ and $\vee$ are used for AND logic and OR logic gates, respectively. The structure in Figure 5(c) has two pipeline stages; (i) AND stage (P-block), (ii) OR stage (N-block). It is formed by cascade connection of the stage containing six AND gates (P-blocks) followed by the other stage containing an OR gate (N-block). In Figure 5(d), the pipeline structure has three stages AND-OR-AND. In a single clock cycle, logic signal traverses two stages namely P-blocks and N-blocks. AND gates do evaluation during low level of the clock and OR gates evaluate during high level of the clock signal in the system. If the two TSPC implementations are compared, 3-stage implementation operates at higher clock rates than 2-stage implementation, because the logic gates have less number of gate inputs, that is less number of parallel transistors at the gate inputs. On the other hand, initial delay in 3-stage implementation is 1.5 clock period, however, for 2-stage implementation the initial delay is only one clock period.

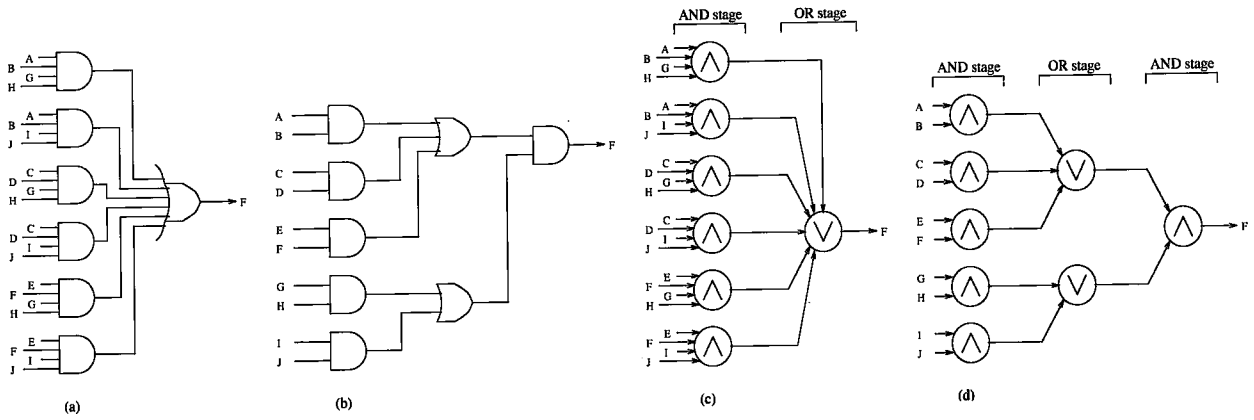
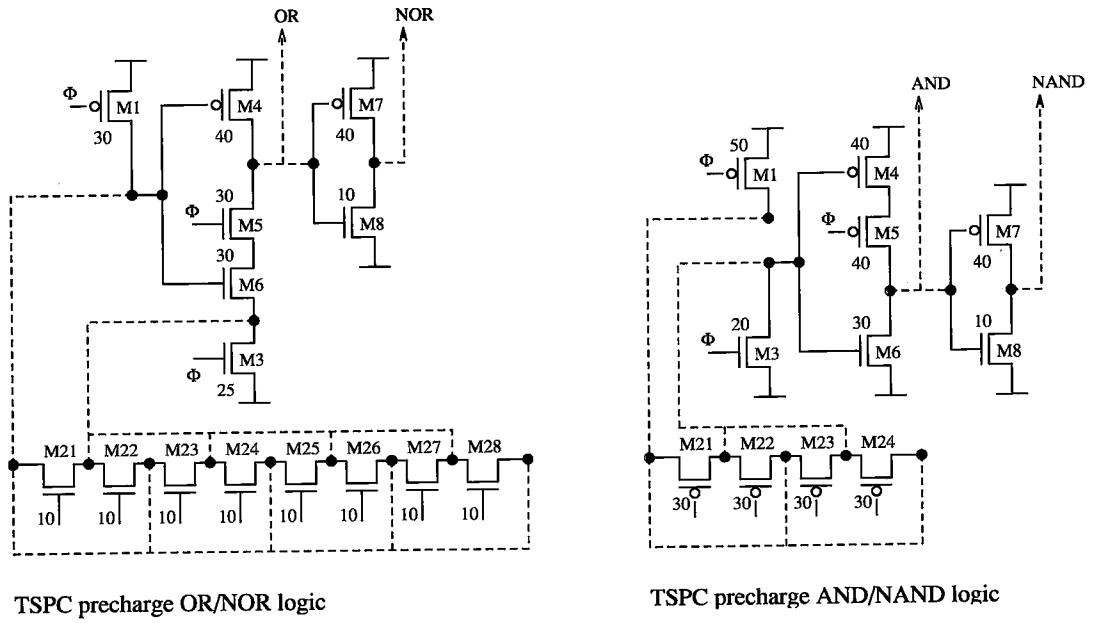


Figure 5. Boolean function, F, implemented by (a) 2-stage static, (b) 3-stage static, (c) 2-stage TSPC pipelined, (d) 3-stage TSPC pipelined AND and OR logic gates.

3. Core Cells of Pipeline Structured Gate Array

The new pipeline structured gate array architecture (PSGA) consists of two kinds of core cells which are AND/NAND and OR/NOR logic gates. These core cells are designed in TSPC circuit technique with the described logic style above and have a single logic depth. The NAND gate and the NOR gate are implemented by connecting static inverter to the outputs of AND and OR gates, respectively.



TSPC precharge OR/NOR logic

TSPC precharge AND/NAND logic

Figure 6. The TSPC circuit schematics of core cells; OR/NOR and AND/NAND.

The core cells are designed in 3μm CMOS n-well technology [8]. AND/NAND gates with up to 4 inputs and OR/NOR gates with up to 8 inputs are designed. As a design specification, 200 MHz clock

frequency is chosen for 5 volts supply, fan-out of 8 and 2 for AND/NAND and OR/NOR gates, respectively, with 1.0 mm gate output wiring length. The circuit schematics of the core cells are given in Figure 6. The dashed lines indicate possible connections in the circuits. These connections are to be routed by metalization during the gate array process. The circuit of an OR gate given before in Figure 2(d) is modified as suggested in [7]. A delay is added by M3 to the falling signal at the drain of M4 in order to wait for the signal to settle down at the drain of M1. Thus, undesirable fall and rise (dips) of the signal at the drain of M4 is avoided. The result is faster OR/NOR gates. The modification is not applied to the AND/NAND gates because the p-transistors connected in series degrade the speed of the circuits more severely than n-transistors.

The transistor sizes are settled after a number of iterative SPICE simulations. The channel length of all transistors is 3μ and the width of each transistor is given next to transistor in Figure 6. Using the design rules in [8], each input of OR/NOR gate and AND/NAND gate has 24fF and 72fF gate capacitance, respectively. The delay characteristics of the core cells are given in Figure 7, which shows the delay versus number of inputs for different number of fan-outs. For NAND gate with fan-out 8, gate delay is 2.6 nsec and for a NOR gate with fan-out 2, gate delay is 2.4 nsec. Therefore, the clock period is 5.0 nsec, hence the clock frequency is 200 MHz as specified. Although more than 200 MHz can be obtained by smaller fan-outs and gate inputs, clock rise and fall times below 1.0 nsec are hard to achieve using CMOS clock drivers. Clock driver design is kept out of the scope of this paper, but for a number of gates with 250fF clock load per gate, 1.0 nsec rise and fall times, which are 20 % of the 5.0 nsec clock period, can be easily achieved in CMOS.

From the simulations, the power dissipation of the gates is found out to be mainly due to the precharge and pre-discharge intervals. Therefore even if the gate input does not change state, power is consumed by the gates and it varies very slightly as the number of gate inputs and the fan-out vary. Worst case power dissipation at 5 volts supply is 28 μW/MHz for a 4-input AND/NAND gate with fan-out 8 and 14 μW/MHz for 8-input OR/NOR gate with fan-out 2 where 1.0 mm wiring length at the output is assumed in the simulations.

4. Gate Array Architecture

The pipeline structured gate array (PSGA) architecture is composed of alternating rows of P-blocks and N-blocks as masters separated by routing channels. The personalization of the masters is done on contact, metal-1, via and metal-2 layers. The core cells AND/NAND gates are routed on P-block masters and the core cells OR/NOR gates are routed on N-block masters. The architecture is shown in Figure 8. Parallel connectable transistors of each P-block and N-block are separated from the precharge and latch circuitry to simplify the routing between P-blocks and N-blocks. The layout of two master P-blocks and two master N-blocks is given in Figure 9.

The PSGA architecture is obtained when the layout is vertically and horizontally repeated by abutment. Power supply lines run vertical in metal-2. The widths of these metal-2 wires depend on the number of P-blocks and N-blocks arranged in the vertical direction and the clock frequency. In the routing channels,

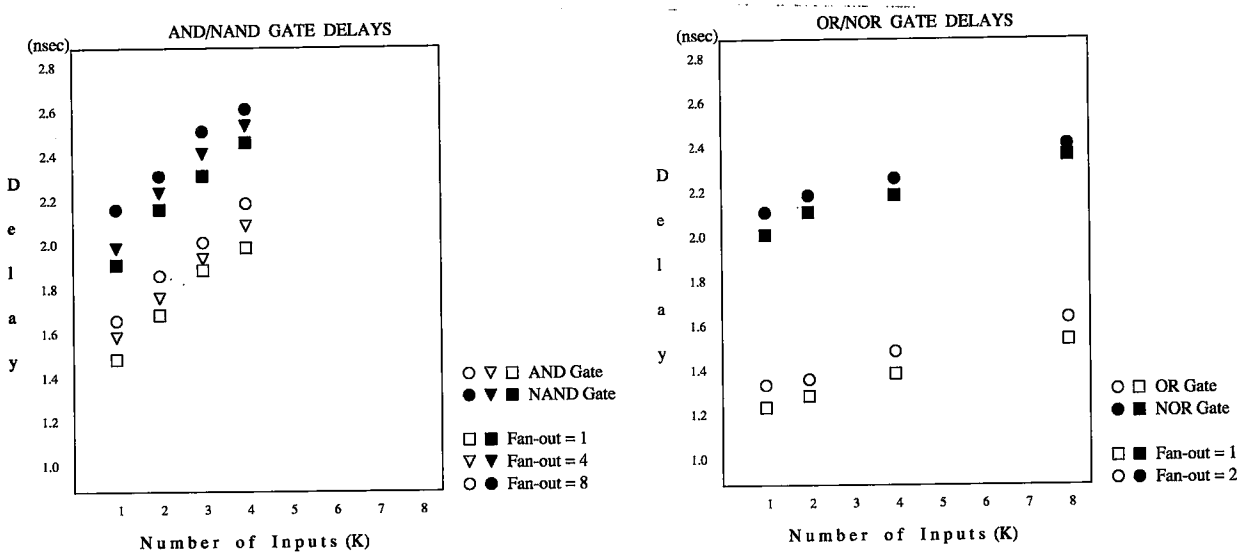


Figure 7. The delay characteristics of AND, NAND, OR and NOR gates.

arranged in the vertical direction and the clock frequency. In the routing channels, basically horizontal wires run in metal-1 and vertical wires run in metal-2 and polysilicon. The direction of routing in the channels for metal-1 and metal-2 lines is not restricted. Overcell routing is also possible for metal-2. Clock lines run vertical in metal-2 for each column. Each N-block has 6 parallel connectable n-transistors which are implemented on a single diffusion island. For 8 input OR/NOR gate realization, 2 n-transistors are borrowed from adjacent N-block. Each P-block has 4 parallel connectable p-transistors on a single diffusion island. The parallel connectable transistors are separated from each other by gate isolation technique in which separator transistor is turned off by connecting its gate to ground for n-transistor and to Vdd for p-transistor.

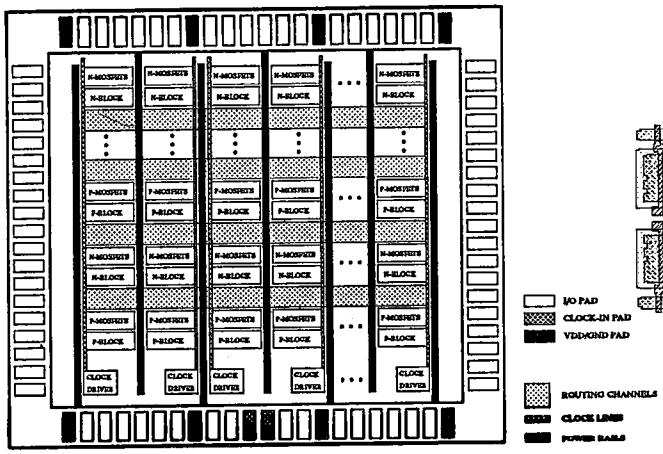


Figure 8. The pipeline structured gate array architecture.

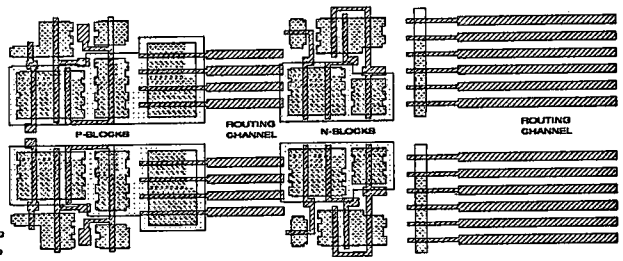


Figure 9. The master P-blocks and N-blocks.

Both P-block and N-block have an n-transistor and a p-transistor with common gate that can be routed as an inverter connected to the outputs to obtain the core cells NAND gate and NOR gate, respectively.

At very high frequencies, clock distribution is a problem even for single phase clock. Multiple clock drivers are used in the architecture to distribute the clock loading. The core cells operate at 200 MHz with 1.0 nsec clock rise and fall times. Therefore, the clock drivers can be Schmitt-trigger type circuits to sharpen the rising and falling edges of the clock. The clock driver's load for each gate is about 250 fF. This much of capacitance requires 6 μW/gate/MHz power at 5 volts supply. The clock signal propagation direction is the opposite of the data signal propagation direction, so that the system operates in safe [7]. The core cell characteristics of the proposed layout style are summarized in Table 1.

Table 1. The characteristics of the core cells

Technology	3.0μm CMOS
Frequency	200 MHz
Power	4.2 mW/gate @ 200 MHz
NMOS/PMOS Count	15/11
AND/NAND gate area	119 μ $\times$90 μ
OR/NOR gate area	90μ $\times$94 μ

The proposed logic style is compared with other possible logic styles of TSPC circuit technique. The OR/NOR and AND/NAND gates are realized using the TSPC dynamic logic circuits given in Figure 1(b) and (c). To obtain similarity with the proposed logic style, an 8-input OR/NOR gate is realized for the N-blocks, and a 4-input AND/NAND gate is realized for the P-blocks given in Figure 1(b) and (c). In Figure 1(b), P-logic is the complementary of N-logic and in Figure 1(c), static logic is well known complementary CMOS. The gates are simulated using SPICE and the results are tabulated in Table 2. The complementary part of the parallel connected transistors contains transistors connected in series, therefore as a disadvantage, the delay of the gate becomes very large and the clock frequency decreases. The area of the gate increases also due to the complementary side of the circuits. On the other hand , the advantage of the complementary circuit is seen as decreased power dissipation. However, 200 MHz clock frequency could not be reached in other logic styles of the TSPC circuit technique.

Table 2. Comparison among different logic styles in TSPC circuit technique

	Proposed			Figure 1(b)			Figure 1(c)		
Gate[1]	Size[2]	F^3_{CLK}	Pd[4]	Size[2]	F^3_{CLK}	Pd[4]	Size[2]	F^3_{CLK}	Pd[4]
8-input OR	.0085	200	2.58	.014	77	1.26	.016	62	1.98
8-input NOR	.0085	200	2.78	.014	77	1.55	.016	62	1.4
4-input AND	.01	200	5.1	.013	77	1.69	.016	62	2.4
4-input NAND	.01	200	5.5	.013	77	2.36	.016	62	1.83

1. OR/NOR have fan-out of 2, AND/NAND have fan-out of 8.
2. Gate area is in square mm.
3. The maximum clock frequency for the slowest gate, common to all gates. Duty cycle is 50 % with 1.0 nsec rise/fall times.
4. Power dissipation for the worst case input sequence at given maximum clock frequency and at 5 volts supply.

5. Design Examples

i. An OR-NAND and NOR-AND Example:

The logic diagram of the logic functions O and P to be implemented on the PSGA architecture as an example is given in Figure 10(a). The functions O and P are designed using the core cells and the corresponding pipeline structure is shown in Figure 10(b). When the logic diagram is implemented on the pipeline structure, timing becomes important. The inputs to the first stage should be valid before the rising edge of the clock and the inputs to the second stage should be valid before the falling edge of the clock. The layout of the logic function is given in Figure 11.

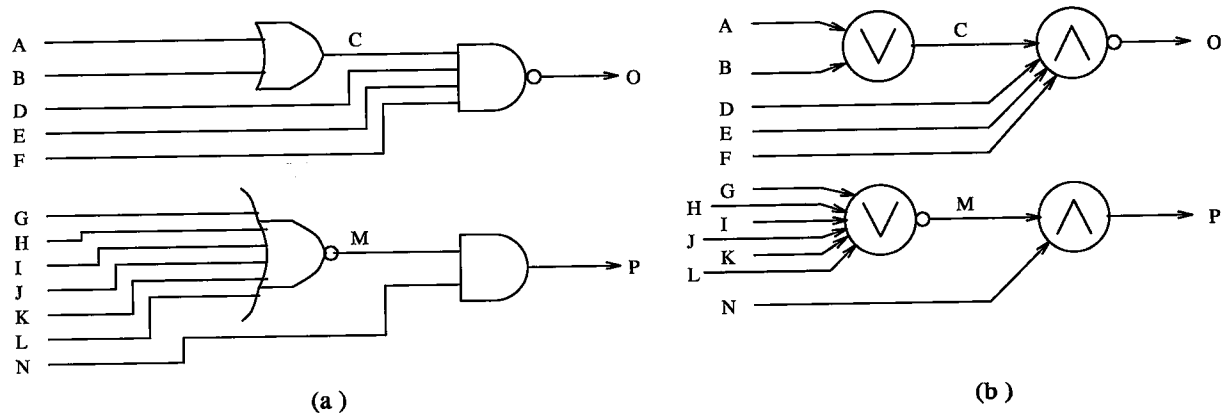


Figure 10. (a) Logic diagram of OR-NAND and NOR-AND gates. (b) The same logic is designed using the core cells in the pipeline structure.

ii. A Serial Full Adder Example:

A serial ful adder is also implemented as a design example on the PSGA. The Boolean functions of the sum output *Sum* and the carry output *Cout* are expressed as,

$$Sum = \overline{(\overline{E} + B)(E + \overline{B})}$$

$$Cout = \overline{(\overline{E} + \overline{B})(E + \overline{Cin})}$$

where A, B, and Cin are the inputs to the full adder and

$$\overline{E} = (\overline{A} + Cin)(A + \overline{Cin})$$

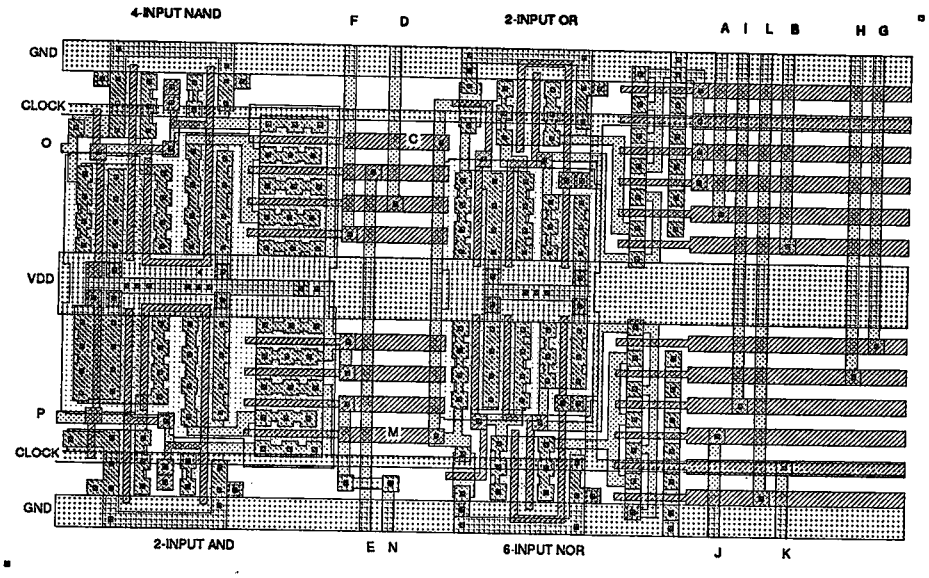


Figure 11. The logic function implemented by OR-NAND and NOR-AND gates is routed on the masters of the pipeline structured gate array.

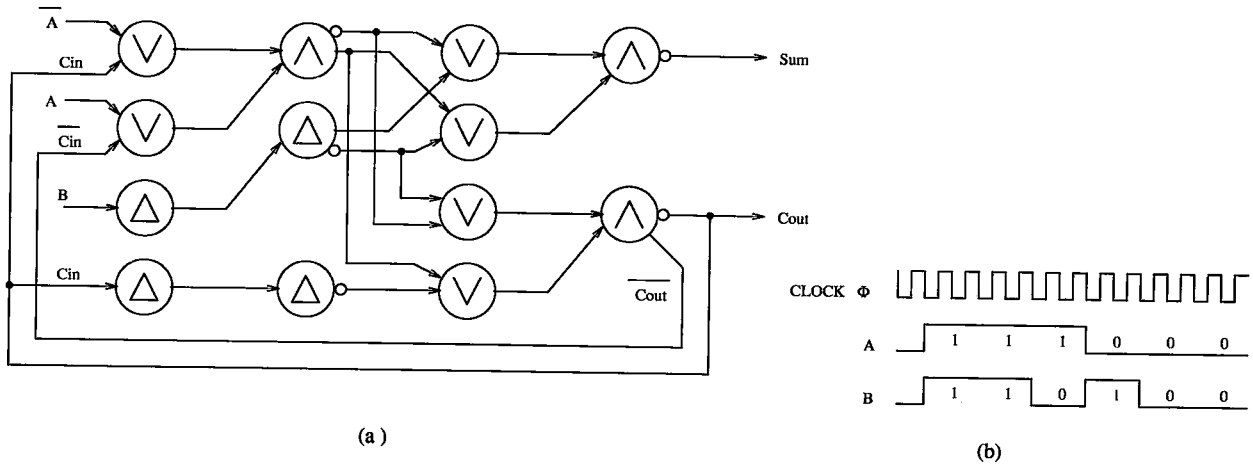


Figure 12. A serial adder design using the core cells of the pipeline structured gate array.

The adder is implemented in four pipeline stages. Connections of the cells in the adder are drawn in Figure 12(a). The symbol Δ represents a delay element in the figure. The delay element inserts one half the clock period delay to the signal flowing through it. It can serve both as inverting and non-inverting delay element simultaneously. Two non-inverting, one inverting and two simultaneously inverting and non-inverting delay elements are used in the serial full adder. In TSPC P-block, the delay element is implemented by a single input AND/NAND gate and it is implemented by a single input OR/NOR gate in TSPC N-block. The serial full adder is simulated using SPICE. Input waveforms for the simulation are drawn in

Figure 12(b). *Cout* and its inverted signal are cleared after applying a number of zeroes to the inputs A and B. The given input waveforms correspond to A $=(0111)_2$ and B$=(1011)_2$. The result in serial form (LSB first) is $(10010)_2$ from *Sum* output. For each input value, *Cout* is calculated after two clock cycles, therefore the next significant bit at the input can only be applied after two clock cycles. Similarly, the *Sum* output is evaluated in every two clock cycles. The initial delay is two clock periods and the throughput is at half the clock frequency. The SPICE simulation results for *Sum* and *Cout* outputs are shown in Figure 13. The *Sum* output shows the result of addition as $(010010)_2$ as expected. The clock frequency is 200 MHz and the power dissipation is 23.5 mW.

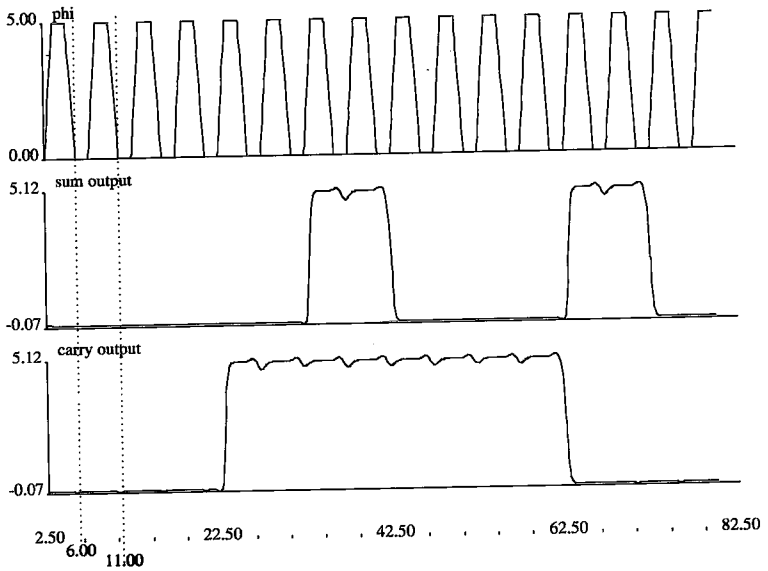


Figure 13. Serial adder SPICE simulation result.

6. Conclusion

In ASIC designs, there has always been a demand for gate arrays for many design applications which do not require very high IC performance. Designs applicable to the pipeline structure that operates at very high clock frequencies are preferred to be implemented by full-custom design methodology which needs very long design time. In this study, a logic style in the TSPC circuit technique is used to implement basic CMOS gates in pipeline structure embedded into a new pipeline structured gate array architecture (PSGA) operating at 200 MHz. It is shown that serial connected transistors, especially p-transistors in a TSPC circuit degrade the speed of the circuit and therefore only parallel connected transistors are chosen as the fastest elementary logic gates with logic depth of one. It is found out that these logic gates can constitute very regular structure which are highly applicable to a gate array architecture. The PSGA is constructed with routing channels instead of channelless architecture in order to reduce parasitic capacitances and resistances which become dominant at very high clock frequencies. The TSPC precharge logic technique is used for the core cells in the gate array. The core cells are composed of AND/NAND and OR/NOR gates and they are implemented by a logic style of parallel connected transistors. Each core cell is full-custom designed to operate at 200

MHz clock frequency. The personalization on the new gate array architecture is done with cut, via, metal-1 and metal-2 layers.

It is demonstrated with the examples that any Boolean function can easily be realized in various ways by using the core cells in the PSGA architecture without any excessive design effort due to device sizing and layout complexity. An OR-NAND NOR-AND design example is presented for the core cell routing demonstration over the PSGA architecture. Another design example, serial full adder is implemented in four pipeline stages and it is verified by SPICE simulation that the adder operates at 200 MHz clock frequency.

The proposed logic style is compared with three other possible logic styles in TSPC circuit technique. It is seen that the proposed logic style operates about three times faster with two times more power dissipation in 1.5 times smaller area than the other logic styles.

The power dissipation of the gates and the clock frequency limits the number of gates in an IC package. 1K gates in about 10 mm^2 operating at 200 MHz dissipate about 4.2 W. One can expect to reach the same frequency for sub-micron CMOS technology and to consume much less power due to smaller internal node capacitances.

The PSGA architecture can find applications in the area of concurrent system designs like high speed signal processing in which the designs can be implemented on pipeline structures [9] or bit-serial and digit-serial arithmetic operations implemented on systolic arrays [10]. Future work for the pipeline structure gate array can be the development of a CAD software which does the selection of the core cells, simulation of the pipeline design and routing of the cells over the gate array.

REFERENCES

[1] M. A. Beunder, J. P. Kernhof and B. Hoefflinger, "The CMOS Gate Forest: An Efficient and Flexible High-Performance ASIC Design Envionment" *IEEE J. Solid-State Circuits,* vol.23, no.2, pp. 387-399, April 1988.

[2] C. A. Mead and L. Conway, *Introduction to VLSI Systems.* Reading, MA: Addison-Wesley 1980.

[3] N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design.* Reading, MA: Addison-Wesley 1985.

[4] Y. Ji-Ren, I. Karlsson and C. Svensson, "A True Single-Phase-Clock Dynamic CMOS Circuit Technique," *IEEE J. Solid-State Circuits,* vol. SC-22, no. 5, pp. 899-901, Oct. 1987.

[5] R. H. Krambeck, C. M. Lee and H. S. Law, "High-Speed Compact Circuits with CMOS" *IEEE J. Solid-State Circuits,* vol.SC-17, no 3,pp. 614-619, June 1982.

[6] N. F. Goncalves and H. J. De Man, "NORA: A Racefree Dynamic CMOS Technique for Pipelined Logic Structures," *IEEE J. Solid-State Circuits,* vol. SC-18, no.3, pp. 261-266, June 1983.

[7] J. Yuan and C. Svensson, "High-Speed CMOS Circuit Technique," *IEEE J. Solid-State Circuits,* vol.24, no.1 pp. 62-70, Feb. 1989.

[8] Invomec MPC Service, *Design Rules, 3 micron double poly, double metal, n-well* CMOS, April 1987

[9] M. Hatamian and G. L. Cash, "Parallel Bit-Level Pipelined VLSI Designs for High-Speed Signal Processing," *Proceedings of the IEEE,* vol. 75, no. 9, pp. 1192-1202, Sep 1987.

[10] P. Corbett and R. Hartley, "Designing Systolic Array Using Digit-Serial Arithmetic," *IEEE Tran. Circuits Syst.-II,* vol. 39, no.1, pp. 62-64, Jan. 1992.