Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Analysis of Delay Limitation and Circuit Power Balance Optimisation for CMOS Based Circuits

Ruotong Li *

School of Microelectronics, Tianjin University, Tianjin, 300072, China

* Corresponding Author Email: 3021232021@tju.edu.cn

Abstract. As integrated circuits and technology have advanced, people's requirements for integrated circuits are getting higher and higher, and integrated circuits with high performance, low latency and low power consumption characteristics are needed to satisfy all kinds of human needs. However, meeting these needs requires a balanced optimization of circuit delay and energy consumption. The gate size and voltage need to be varied simultaneously, and the most suitable voltage and gate are found by combining the gate size and voltage ratio using a linear programming solver. This method can find the optimal gate size and voltage for minimum power circuits with delay requirements, and it also provides new possibilities for comprehensive optimization of Complementary Metal Oxide Semiconductor (CMOS) integrated circuits from both performance and energy perspectives and provides new ideas for IC optimization. In the future, it is necessary to continue to explore the automatic optimization techniques and use deep learning techniques to obtain more efficient circuit optimization methods.

Keywords: Integrated Circuit; Power Consumption; Delay; Gate size; Voltage.

1. Introduction

With the rapid development of integrated circuits, nowadays, integrated circuits have penetrated all corners of the society and have a very wide range of applications in medical, communication, and transport. In the pursuit of higher performance, more complex functions and smaller size at the same time, the power consumption and propagation delay of the circuit will rise at the same time, from the point of view of circuit efficiency and reliability as well as saving energy and packaging costs, the optimization of the circuit's propagation delay and power consumption has become an urgent issue [1]. CMOS is the most used logic style for integrated circuits, and the optimization of CMOS circuits has become the key to solving this problem.

In this paper, the basic structure of the inverter, which is the fundamental unit of CMOS integrated circuits, and its working principle are introduced, from which the concepts of propagation delay and circuit power consumption are introduced, and the subsequent chapters explain why delay optimization is necessary, how to optimize the delay, why power optimization is necessary, what constitutes power optimization, and how to optimize the power consumption, and finally, the power optimization methodology is presented in the form of a 4-bit absolute value detector as an example, a power optimization method based on a 1.5 times minimum latency circuit is proposed - gate sizing and voltage scaling are combined to find the most suitable voltage and gate by using a linear programmer solver.

In this article, CMOS circuits are briefly introduced, and how circuits can be optimized is analyzed from the perspectives of delay and power consumption, and a circuit power optimization method based on constrained delay is proposed, which is able to minimize the power consumption in case of meeting the performance requirements, and provides new possibilities for the optimization of CMOS integrated circuits from the perspectives of both performance and energy in an integrated manner.

2. Basic Theoretical knowledge

2.1. Concepts of CMOS Integrated Circuits

As the basic unit of CMOS logic circuit, the inverter has a structure as shown in the figure 1 below.

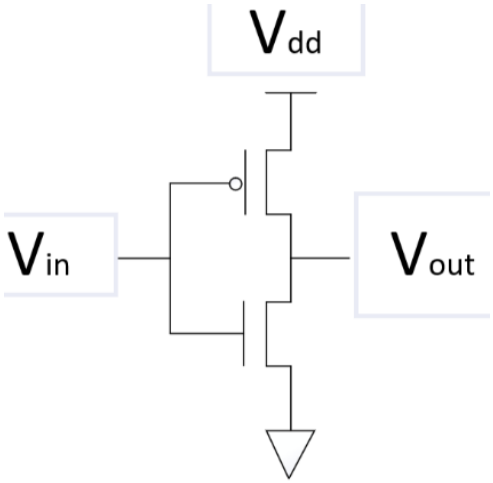

Figure 1. Structure of an inverter (Photo/Picture credit: Original)

A CMOS inverter is composed of an NMOS and a PMOS. The gate of the PMOS is connected to the gate of the NMOS, which is regarded as the input terminal of the inverter. The drain of the PMOS is connected to the drain of the NMOS, which is called the output terminal of the inverter. Whereas the source and substrate of NMOS are connected to ground, the source and substrate of PMOS are connected to the power supply terminal, or Vdd. When the input is low level (Vin = 0V), the output terminal is high level (Vout=Vdd). When the input voltage gradually increases so that the gate voltage is equal to the voltage supplied, the output terminal will be discharged to zero potential through the n-channel. This is what an inverter does.

As a widely used logic style, the main feature of CMOS is its low power consumption. The CMOS circuit has a special complementary structure of field effect transistors. During operation, one of the two field effect transistors is stationary, and the other is conducting, so the static power consumption is nearly equal to zero in theory. However, due to the static power consumption caused by the existence of leakage current losses in the transmission process, and the dynamic power consumption that includes three types of power consumption, namely, capacitive power consumption caused by logic jumps, competing risky power consumption caused by pathway delays, and short-circuit power consumption caused by instantaneous conduction of circuits, there is a small amount of power consumption in the CMOS integrated circuits [2, 3].

2.2. Propagation Delay

The time which takes for digital signals to move from the location in a logic circuit to another is regarded as the propagation delay. Reducing propagation delay is crucial in digital circuits to guarantee that signals reach their destinations on schedule. In VLSI design, many techniques are used at each design stage to minimize propagation delay. For example, buffer insertion, using fast logic families, transistors sizing, negative group delay (NGD) active circuit, inserting high-k dielectric materials are used [4, 5]. Methods usually used is detecting the critical path and then optimize the circuit to reduce the critical path delay [6]. However, limiting propagation delay also involves trade-offs. It is known that a VLSI computation that implements a specific algorithm needs more energy the faster it computes, so techniques like buffer insertion and using fast logic families would increase power consumption.

Although reducing propagation delay is critical, designers must weigh its advantages against possible trade-offs. This is an important problem which needs to be solved.

2.3. Circuit Power Consumption

In integrated circuit design, another issue that needs to be considered is power consumption. Power consumption is an important factor in ensuring the function of circuit. Power consumption is related to performance, the higher the power consumption, the better the performance. However, as the level of integration of integrated circuits increases, the greater the power consumption can lead to high circuit temperatures or even circuit damage, while excessive power consumption can also cause difficulties in packaging [1]. As a result, people should focus on energy consumption when designing circuits.

As mentioned in section 2.1, static power consumption and dynamic power consumption constitute the power consumption in CMOS circuits, and hence the power consumption can be optimized from these two perspectives. From this aspect, design techniques used in CMOS integrated circuit with low power consumption can be separated into three parts: dynamic power design, static power design, and low-power integrated technology [7]. More detailed methods will be described in later chapters.

Transistor sizing plays an essential role in circuit design when considering the purpose of increasing the speed and reducing the power consumption of integrated circuits [8].

3. Optimization of Delay and Circuit Power Consumption Analysis

3.1. Optimization of Propagation Delay

Propagation delay is the time which takes for signals to move from the input to the output of a CMOS gate. There are two parts to propagation delay: extrinsic delay and intrinsic delay. Extrinsic delay is the result of external elements like connection capacitance and wire resistance, whereas intrinsic delay is the result of the internal design and operation of the gate. Problem can be solved from these aspects.

To reduce propagation delay, tons of methods have been proposed. Many scholars posted that transistor sizing was a good way to improve the delay, however, this way would have a significant effect layout area exist. In 1995, a method named transistor reordering was proposed to complete delay optimization with little effect on layout area exist, this method targets the reduction of charging/discharging events in the logic gate’s internal capacitance to optimize delay [9]. Rearranging the transistors in a MOS logic gate may be a useful method of reducing latency because the transistor order can have a big impact on the propagation delay of the gate [10]. This strategy was improved in 2022, and the suggested algorithm reorders by taking into account both the transistor's input characteristics and its likelihood of being active [11].

Besides, another technique to reduce delay is called variable input delay (VID) technique. VID is a way of repairing clock signals. Much of the dynamic power consumption is attributed to clock delay. Square-wave signals must be timed simultaneously by multiple modules in highly synchronous circuits. As a result, there is significant power consumption due to the wasted clock signals that are used by various modules. As a result, by supplying clock signals to idle gates, this method avoids the delay [12]. In this article, the author used the concept of VID to design circuits for both two-input NAND gates and full adder. The gate delay, net delay and total delay of the circuit designed using this method are all much smaller than circuits designed without VID technique. It is clear that VID is a good method for circuit delay optimization.

The last technology talked about optimizing delay is buffer insertion. The most influential technique posted by Van Ginneken gave the possible locations of buffers and helped solve the delay optimization problem effectively [13]. Buffer insertion is thought to be highly effective at reducing propagation latency, to increase timing performance, additional buffers are required as the design size decreases. However, the buffer consumes energy, which means that the more buffers are inserted, the more energy is consumed. That is a tricky problem need to settle.

3.2. Circuit Power Consumption Optimization

Since integrated circuit power consumption has two strategies: static power consumption and dynamic power consumption, these two aspects are also considered in the design of circuit power consumption optimization. Leakage current primarily causes the static power consumption, which is mainly related to process control, power supply voltage and threshold voltage. Dynamic power consumption consists of three parts. As mentioned in Section 2.3, it can be optimized from five levels: physical level, logical level, RTL level, behavioral level and system level, and the optimization effect decreases step by step.

In the past, the main part of the circuit power consumption is the dynamic power consumption, but with the reduction of feature size, leakage current is getting larger and larger, and static power consumption has become the main part of circuit power consumption optimization [7].

Different methods exist for reducing leakage power depending on how a system operates. Many techniques, including as technology scaling, voltage scaling, clock frequency scaling, switching activity reduction, etc., were employed to reduce static power.

Multi-Vth optimization is a very useful technique to reduce leakage currents to optimize static power [14]. To maximize power and delay, this method makes use of transistors with numerous threshold voltages (Vth). Here, clock durations were kept to a minimum by using low voltage components in crucial delay paths. To minimize static leakage power on non-critical channels without causing a delay cost, higher voltage devices were employed [15].

As for the optimization of dynamic power consumption, many techniques are used at different levels. For transistor level, high-k materials are used for gate dielectric [16]. By using high-k materials, the gate leakage is reduced effectively, although it will increase short-channel effects. For circuit level, transistor sizing can be used. An integrated circuits transistor size has effect on both the energy loss and the gate delay [16]. Transistors with wider sizes in logic gates will have smaller gate latency, but the gate's switching energy consumption will rise. To cut consumption without changing the critical path delay, calculating the positive slack is a useful method.

3.3. Circuit Power Optimization Based on Delay Limiting

In actual integrated circuit design, it is impossible to satisfy only a single optimal delay or optimal power consumption. The most important thing is to find a balance between the two. When optimizing power consumption for a circuit with delay constraints, it is necessary to consume as little energy as possible while meeting the actual required delay conditions.

Taking a 4-bit absolute value detector as an example, energy loss needs to be minimal when the transmission delay is equal to 1.5 times of minimum latency. Sizing the gate and supply voltage scaling should be both used.

Regarding the computation and optimization of the delay, it is necessary to apply the theory of logical effort [17]. The sum of the delays for each stage yields the overall delay.

$$Delay = D \times t_{p0} \tag{1}$$

For this formula, tp0 is named parasitic delay of an inverter and the formula for solving D will be given below.

$$D = g \cdot h + p \tag{2}$$

$$h = \frac{C_{out}}{C_{in}} \tag{3}$$

Besides, the impact of VDD on delay should also be considered, delay can be assumed that it is proportional to VDD as

$$V_{DD} \sim \frac{V_{DD}}{(V_{DD}-V_T)^2} \tag{4}$$

This function satisfies the minimum value when VDD equals 1, hence the delay is minimal. VDD has an impact on energy usage as well. The study makes a 50% delay increase as a compromise. This indicates that when the function equals 1.5 times of the minimum latency, the VDD must be solved. The function value that determines the minimum delay when VDD equals 1V. In conclusion, when examining the effect of VDD on the overall delay, VDD can be equivalent to 0.775.

By using these formulas, it can be easy to find the best size and voltage to meet the requirement of delay. As for the energy consumption, there are some formulas.

$$P_{static} = I_{static} \times V_{DD} \tag{5}$$

$$P_{dynamic} = \alpha \cdot f \cdot {V_{DD}}^2 \cdot (C_L + C) \tag{6}$$

CL represents the load capacitance, C denotes the CMOS structure's internal capacitance, f signifies the operation frequency, and α signifies the transition rate. The internal capacitance is determined during the delay optimization step, and the load capacitance should be designed by the designer. Alpha has a connection to transistor.

After knowing about all the parameters, the optimization of delay and energy consumption sees easy than before. By using a linear programming solver, the best transistor size and supply voltage will be found.

4. Conclusion

It can be known from other scholars' papers, the power consumption increases while the delay decreases as VDD increases. Consequently, in order to maximize the circuit, a balance between power and delay must be achieved. Because in a conventional cell design, dynamic power consumption continues to be the main source of energy consumption.

For the future, it is need to optimize the energy consumption model, making it possible to calculate the dynamic energy consumption more accurately. At the same time, we need to continue to explore automated optimization techniques in depth, combining circuit optimization with deep learning, and thus obtaining efficient circuit optimization methods to reduce the delay and power consumption. UT Austin, in collaboration with Intel, has proposed a method for gate sizing using deep neural networks. The authors trained two deep learning models at the same time, where the critical-network is responsible for evaluating the effectiveness of each gate sizing, and the actor-network is responsible for the most effective sizing method, using this method can significantly reduce the number of searches needed and obtain less design time, but this method is still supervised learning rather than reinforcement learning. not reinforcement learning and applying reinforcement learning methods is still a major challenge.

References

[1] LAN X Z. Design Methodology for Low-Power CMOS Integrated Circuits. Microelectronics, 2004.

[2] Yin Li, Cao Weigang, Wang Xiaohuan. Application of CMOS integrated circuits. Wireless Internet Technology, 2012 (7): 183-183.

[3] Xu YJ. Power consumption estimation and low-power design of integrated circuits. Ph.D. Thesis, Institute of Computing Technology, Chinese Academy of Sciences, April, 2006.

[4] Ravelo B, Pérennec A, Le Roy M. Application of negative group delay active circuits to reduce the 50% propagation Delay of RC-line model. 2008 12th IEEE Workshop on Signal Propagation on Interconnects. IEEE, 2008: 1-4.

[5] Xu P, Pan Z. The insertion of high-k dielectric materials in multilayer graphene nanoribbon interconnects for reducing propagation delay and expanding bandwidth. Organic Electronics, 2020, 78: 105607.

[6] Hashimoto M, Onodeva H. Increase in delay uncertainty by performance optimization. ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No. 01CH37196). IEEE, 2001, 5: 379-382.

[7] Qian F W. Design on CMOS integrated circuit of low power consumption. Proceedings of the 2018 7th International Conference on Advanced Materials and Computer Science (ICAMCS 2018), Dalian, China. 2018: 21-22.

[8] Visweswariah. Optimization techniques for high-performance digital circuits. 1997 Proceedings of IEEE International Conference on Computer Aided Design (ICCAD). IEEE, 1997: 198-207.

[9] Carlson B S, Lee S J. Delay optimization of digital CMOS VLSI circuits by transistor reordering. IEEE transactions on computer-aided design of integrated circuits and systems, 1995, 14(10): 1183-1192.

[10] Carlson B S, Chen C Y R. Effects of transistor reordering on the performance of MOS digital circuits. [1992] Proceedings of the 35th Midwest Symposium on Circuits and Systems. IEEE, 1992: 121-124.

[11] Muñoz M M, Kessler H, Porto M, et al. Transistor Reordering for Electrical Improvement in CMOS Complex Gates. 2022 35th SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI). IEEE, 2022: 1-6.

[12] Bansal P, Somwanshi D K, Dhaliwal B S, et al. The roles of delay and power optimization techniques in VLSI design. Advances in Image and Data Processing using VLSI Design, Volume 1: Smart vision systems. Bristol, UK: IOP Publishing, 2021: 4-1-4-21.

[13] Van Ginneken L P P P. Buffer placement in distributed RC-tree networks for minimal Elmore delay. 1990 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 1990: 865-868.

[14] Sreenivasulu P, Khan P K, Rao K S, et al. Power Scaling in CMOS Circuits by Dual Threshold Voltage Technique. International Journal of Engineering and Innovative Technology (IJEIT), 2013, 3(1).

[15] Suguna T, Rani M J. Survey on power optimization techniques for low power VLSI circuit in deep submicron technology. International Journal of VLSI design and Communication Systems, 2018, 9(1): 1-15.

[16] Raut K J, Chitre A V, Deshmukh M S, et al. Low power VLSI design techniques: A review. Journal of University of Shanghai for Science and Technology, 2021, 23(11): 172-183.

[17] Sutherland I, Sproull R F, Harris D. Logical effort: designing fast CMOS circuits. Morgan Kaufmann, 1999.