Study on the evolution of silicon integrated circuit technology: from scaling to lower power consumption

Yiwen Wu, Ruofei Jia, Ziyang An, Baozhi Ding

Abstract

Silicon-integrated circuit technology has a profound influence on people's daily lives. Throughout history, the goal of the evolution of this integrated technology has changed from scaling to lower power consumption because of multiple difficulties scientists face. By measuring and estimating the trade-off between circuit size and power, scientists have already developed multiple new technologies to upgrade the silicon-integrated circuit. Some new methods people are trying to find out that could be better to solve the problem. We did the research through one of the methods called post-scaling technology. We found out it is possible to control the heat and power consumption of the circuit, which, under the circumstance of minor scaling and stable operation mode.

Keywords: transistor. power consumption. silicon integrated circuit. technology development.

1. Introduction

Transistor is a solid semiconductor device with some functions including detecting, rectification, amplifying, switching, voltage stabilization, signal modulation and so on. As a variable current switch, the transistor can control the output current based on the input voltage.

In the fall of 1945, Bell Laboratory formally established a semiconductor research group including Bratton, Shockley, Bardeen, and others, and began to conduct research on some new materials such as silicon and germanium. In 1947 Bell Laboratory published the first point contact transistor made of germanium semiconductor.

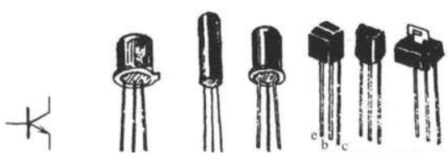

Figure 1 some kinds of transistors 1

For the advantages compared with vacuum tube:

(1) No consumption of components

The tube will deteriorate gradually because of the change of the cathode atoms and chronic air leakage. For some technical reasons, transistors started out with the same issue. With the advances in materials and other improvements, transistors typically last 100 to 1,000 times longer than the electronic tubes.

(2) Less power consumption:

It does not need to heat a filament to produce free electrons, as a tube does. A transistor radio with only a few dry batteries can be listened to for half a year, which is difficult for a valve radio.

(3) Without preheating:

It can run as soon as it's turned on. For example, a transistor radio will make a sound as soon as it is turned on, and a transistor television will appear as soon as it is turned on. Valve devices cannot do this. After the machine is turned on, wait for several seconds to hear the sound, and see the picture. Obviously, in military, measurement, recording and other aspects, transistors are very useful.

(4) Stronger and more reliable:

100 times more reliable than the tube, shock resistance, vibration resistance, which is unmatched by the tube.

The evolution of silicon integrated circuit:

Integrated circuit is that a certain number of common electronic components, such as resistance, capacitance, transistors, and so on, and the connection between these components, through the semiconductor process integrated together with a specific function of the circuit. It is through oxidation, lithography, diffusion, epitaxy, steam aluminum and other semiconductor manufacturing processes, the composition of a certain function of the circuit of the semiconductor, resistance, capacitance and other components and their connecting wires are all integrated in a small piece of silicon, and then welded into a shell of electronic devices.[1]

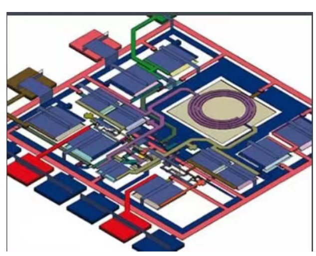

Figure 2 silicon integrated circuit 1

Jack Kilby invented the silicon integrated circuit in 1958. During his work, he came up with the idea that resistors and capacitors could be made of the same materials as transistors. In addition, since all components can be made from the same material, they can be made locally from the same material and then connected to each other to form a complete circuit, and finally, the first integrated circuit was born.[2]

Figure 3 Jack Kilby

For the advantages compared with transistors:

(1) Lower failure rate: The failure rate of integrated circuits is lower than discrete component circuits such as transistors

(2) Simpler circuit: Due to the use of integrated circuit, simplified circuit design, debugging and installation, especially after the use of some special integrated circuit, the circuit is simpler.

(3) Less energy and power consumption: Integrated circuit has the advantages of less energy consumption, the use of integrated circuit than the use of separated electronic components of the circuit power consumption is much smaller.

(4) high cost-performance: Compared with the circuit of discrete components, the circuit composed of integrated circuits has higher performance indicators, and the cost and price of integrated circuits are lower.

(5) Higher reliability: After the adoption of integrated circuit, the welding spot in the circuit is greatly reduced, the possibility of virtual welding is reduced, so that the work of the circuit is more reliable.

2. Why power matters more today than 30 years ago

There are many transistors on an integrated circuit, so the size of the transistor must be small [3], which results in a small voltage applied. In addition, as the distance between applied voltages becomes smaller, the electric field becomes larger, which leads to poor transistor performance. Therefore, we need to apply a very small voltage. When the voltage is very low, the subthreshold swing is not lower. With a transistor that small, the voltage spans from zero to maybe one volt. But when the transistor is large, the voltage can span from zero to maybe ten volts. With today's transistors, we need to reduce the voltage to a very small or even a million times that, and that's just not possible. Because there's a baseline of how low the voltage can go, and that low voltage is thermal fluctuation at room temperature, and there's a natural voltage fluctuation, which is 26 million electron volts. When you apply a voltage, it will fluctuate randomly up and down at that voltage, plus 20 million electron volts, or minus 20 million electron volts, but you can't get the voltage down to that level. It turns out that when the voltage is high, it still gets low, but the subthreshold swing doesn't change, and then the current difference between the transistor on to the transistor off must decrease, and once it decreases, then when the transistor is on, it doesn't have enough current, or you turn it off, it still doesn't get turned off, it's still running. This has led to the current focus on integrated circuits to reduce power consumption [4].

The first part of the report is the reason of why everyone wants to have less power consumed while using the circuit and its benefit.

First of all, due to the functional relationship, part of the power will inevitably be consumed as heat, which will heat up the transistor and reduce the reliability of the transistor. Specifically, first, it will increase the probability of avalanche breakdown of transistors; Second, mechanical failure is caused by different thermal expansion coefficients of different materials; Third, it will cause thermal fatigue of materials.

Secondly, power consumption will bring many hazards. In order to solve these hazards, we must formulate some countermeasures, and the implementation of these measures will inevitably bring many costs. For example, the cooling cost of the chip and the cost of reducing the temperature of the experimental environment.

For the second part of this question, things people have done for the purpose of lower power consumption because power consumption will heat up the whole world, power consumption must be reduced for each chip. This is the formula of circuit power consumption. The first two terms of this formula are dynamic power consumption, and the third term is static power consumption. Low power design technology actually reduces power consumption by changing these parameters. So we only need to optimize dynamic power consumption and static power consumption [5].

$$P = 0.5CU_{DD}^{2}fE_{SW} + Q_{SC}U_{DD}fE_{SW} + I_{leak}U_{DD}$$

To begin with, it is dynamic power optimization technology. Because different design levels have different characteristics in the hardware design process, the dynamic power optimization space at different levels is also different.

The first is the system level. The idea of this level is to perform dynamic power management, that is, according to the change of system workload, selectively set certain system resources to low power mode and complete system tasks with the least component design load or complete the low power mode with the least component design load to reduce the energy usage amount of the system. Currently, energy management strategies can be roughly divided into three types: timeout strategy, prediction strategy and random strategy [6].

The second is the structure and algorithm level. the structure and hierarchy of the algorithm. It selects the corresponding algorithm expression from the circuit structure and coding, reduces the frequency hopping. Common methods include employing structures and parallel lines to optimize coding styles.

Parallelization is to separate the path work into two paths, so the working frequency of each is half. The slow allowed data path is double. Currently, the circuit adopt smaller working pressure. The parallel circuit structure tool is to increase the circuit area based on maintaining circuit permeability to accomplish the goal of doubling energy usage.

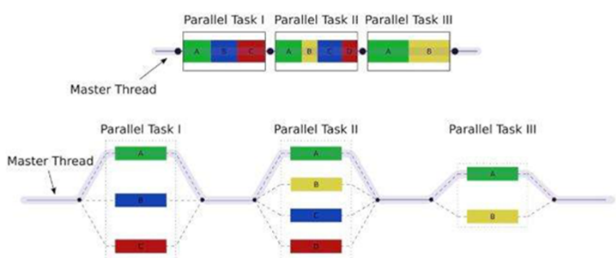

Figure 4 Schematic diagram of parallel structure

The main characteristics of bus in computer system are large load and long lead length, therefore, the bus generally has a big capacity, and it is almost out of the question to transfer its capacity. Therefore, reducing the bus speed is the best method to avoid the electrical fault of the bus. If the same work is performed to reduce the effective speed on the bus, the way to change the code on the data sent on is to change the code.

In addition, there is register transfer level, which is the most commonly used level in circuit design at present. Its basic idea is to isolate the state of some circuits when they are not working, so as to eliminate the invalid turnover of module level circuits in the process of operation. It mainly includes the methods of finite state machine coding optimization, logic reorganization, reducing pseudo jump and so on

Apart from that, there are logic gate levels, which are mainly to reduce the load capacitance, adjust the transistor size, select logic devices with low power consumption, and optimize the logic structure as much as possible. The general ways about gate level power consumption majorization include path balancing, factorial recombination, buffer insertion, phase allocation, etc. [7].

The transistor level is finally reached. Transistor optimization is mainly to reduce power consumption by comparing area and performance of various circuit structures or circuits, or by controlling the course of product. The most typical transistor optimization methods are threshold pressure control method and power input method.

Then there is static power optimization technology. Because the formation of static power is directly related to its process, the optimization level of static power is mostly closely related to circuit and process

There are four main methods:

(1) The process control method mainly reduces the influence of leakage current by controlling the channel length, oxide thickness, junction depth and other structural parameters of transistors, as well as different channel doping methods, such as insulating silicon.

(2) Threshold voltage method the threshold voltage of transistors plays a decisive role in the size of sub threshold current. Therefore, optimizing static power consumption through the control of threshold voltage is a very effective way among many optimization methods. It is also the most common and widely used method in the industry at present. It includes double threshold method, multi threshold method, variable threshold method, dynamic threshold method and so on.

(3) The input control method uses the characteristic that the leakage current of the circuit is easily affected by the input state to properly control the input of the circuit to reduce the leakage. The input vector control method minimizes the leakage power consumption by controlling the input vector state of the circuit when it is not working or inserts stacked transistors into the high leakage cells in the circuit to reduce the leakage. These methods usually have obvious optimization effects only on small-scale circuits

(4) Other control methods adopt different circuit forms and other control methods, which will have a certain effect on the leakage control of the circuit.

The first part is why smaller transistor increase energy consumption in the circuit.

I want to mention is that since the transistor become smaller in the circuit, the circuit can contain more transistor [8]. Since each transistor in the circuit need the same voltage to work, the total power consumption in the circuit increase. We can refer to the graph:

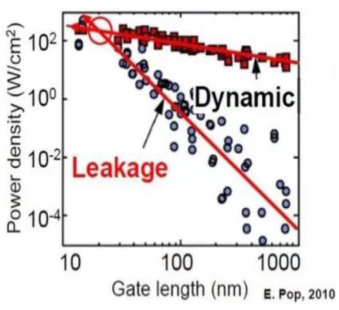

Figure 5 Power-density to Gate Length

The shorter gate length, the greater power density. So in fact, the circuit need more power to work on more smaller transistors.

The second part I want to mention is that how to decrease the energy consumption in the circuit. I generally talk about three ways that we can use to achieve the goal in the circuit.

First of all, the voltage consumption of the equipment can be controlled at the minimum by using advanced technology and proportional reduction technology [9]. Through scaling technology, it can effectively shorten the ratio of transistors, as well as the ratio of interconnecting wires. To achieve the goal of shortening the transistor, we need to reduce the key parameters of the device associated with it and adopt a smaller channel length. After that, we can decrease the energy consumption by controlling the voltage consumption in the equipment.

The second way is encapsulation. We can achieve chip isolation from the outside world through packaging technology, reducing the possibility of external air corrosion to equipment components. We can also use multi-chip packaging, first reduce the related functions of one of the interfaces, to control the delay of the circuit, so as to achieve the purpose of optimizing the circuit. To reduce energy consumption, decreasing air erosion must be a good way to make the circuit more endurable.

The third aspect is to improve the efficiency of switch in the circuit Cutting off the starting resistance, reducing the switching frequency and the switching times can reduce the standby loss can improve the standby efficiency. Specific methods like reduce the clock frequency; Switch from high frequency mode to low frequency mode. To improve standby efficiency, the resistance channel must be cut off after start-up.

If the controller does not have a special starting circuit, it can also be connected to the capacitor in series in the starting resistance, and the loss can be gradually reduced to zero after starting. The disadvantage is that the power supply cannot restart itself. Only after the input voltage is disconnected and the capacitor discharges, can the circuit be started again.

In general, I talk about two topics that why smaller transistors in the circuit increase consumption and how to decrease the energy consumption in the circuit.

While guiding principles always have significant impact, it can also be changed easily due to both internal and external factors.

Nowadays, as technical difficulties associated with microfabrication is increasing. The progress of making new generations of silicon integrated chips has been delayed. In addition, the cost of technological development and production seems to be another big issue. The prize of new chips has been increased because Companies need to keep their amount of benefit through the trade, while they are paying huge amount of money into the development. Then the problem has been kicked to the market. Chips buyers cannot afford such high prize to buy a new chip and the market demand will decrease. Such a chain reaction would finally lead to the change of guiding principle, which has been realized by scientist since a long time ago---- from More Moore to More than Moore. (Figure 6 shows how guiding principles change from past to present)

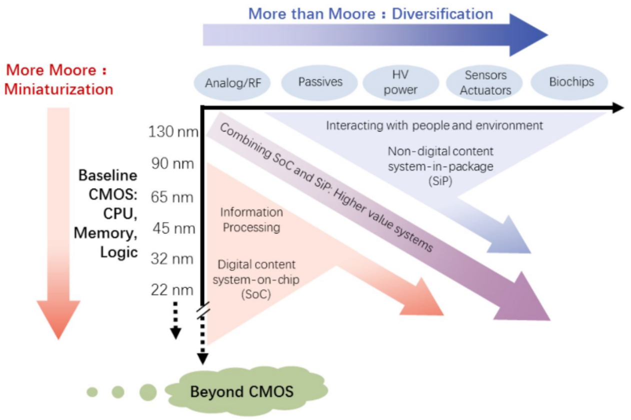

Figure 6 change of guiding principle

As guiding principle has changed, it is important to accelerate the research and development of technologies that can realize high performance, low power consumption, and improved functions and integration using CMOS devices as a base, rather than rely on microfabrication technology alone.

One thing the scientists have been doing to follow the new guiding rule is to develop a new technology called postscaling technology. the postscaling technology is considered as a technology that does not directly rely on the CMOS device scaling. Figure 7 shows the diagram of the postscaling technology. There are multiple technological challenges related to ULSI devices; generally speaking, these challenges can be listed into two basic parts, i.e., increasing of the current driving capability of MOS transistors and increasing of the degree of integration of transistors.

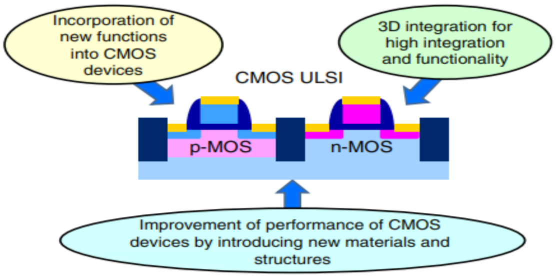

Figure 7 conceptual diagram of postscaling technology with three points explained in the view

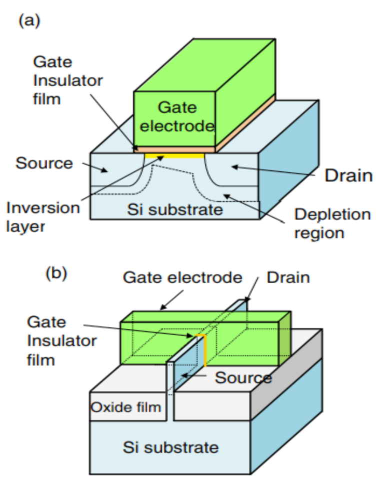

Figure 8 Structure of MOS transistors: (a) planar transistor and (b) FinFET

3. Conclusion

In this review, the evolution of silicon integrated circuit technology and its facing problem has been explained. Under such circumstances, the development of new materials is indispensable to realize devices with high performance, high functionality, and low power consumption independently of downscaling. It will significantly affect the future direction of devices and systems. Referring to the existence research experiences; new technologies should be incorporated together with Si platforms to produce new devices following the new direction. We are looking forward to seeing that the development utilizing materials technology and process technology can lead to the development of new technological applications. In the end, the evolution of the silicon integrated system will continue, with all its benefit to human it will leads to a bright future.

References

[1] Z. Ke, Silicon integrated circuit technology and development trend, 2004, DOI: 10.16257/j.cnki.1681-1070.2004.04.002

[2] Z. G. Yu, Development trend and prospect of silicon integrated circuit, 2003, DOI: 10.16257/j.cnki.1681-1070.2003.02.001

[3] N. Nishiyama; T. Amemiya, On-Silicon Photonic Integrated Circuit toward On-chip Interconnection and Distributed Computing, 2021, DOI: https://ieeexplore.ieee.org/document/9508658

[4] Z. P. Zhou, B. Yin, Q. Z. Deng, X. B. Li, and J. S. Cui, Lowering the energy consumption in silicon photonic devices and systems, 2015, DOI: https://doi.org/10.1364/PRJ.3.000B28

[5] A. V. Bhaskar, "Estimation of Power Consumption in a Network-on-Chip Router," 2022 IEEE Delhi Section Conference (DELCON), 2022, pp. 1-7 DOI:10.1109/DELCON54057.2022.9753477.

[6] C. J. Wei. A review of power optimization technology of integrated circuits[J].Journal of Shanghai Electric Power University,2011,27(2):187-192.
DOI: 10.3969/j.issn.1006-4729.2011.02.020.

[7] Y. Tokusashi, H. Matsutani and H. Amano, "Key-value Store Chip Design for Low Power Consumption," 2019 IEEE Symposium in Low-Power and High-Speed Chips (COOL CHIPS), 2019, pp. 1-3,.
DOI: 10.1109/CoolChips.2019.8721352.

[8] Electronic audiophile forum integrated technology exchange (2020/11/26) https://bbs.elecfans.com/jishu_2012759_1_1.html

[9] EETOP CEO BiJie(2021/12/28) https://zhuanlan.zhihu.com/p/450706717