A seminar paper about

The Fututre of Semiconductors

Aspir Ahmet - 1518066
Stefan Kaufmann - 1518064

02.06.2017

Contents

1 Abstract

The modern semiconductor industry is facing several problems at the moment. In this paper we want to address the principle of semiconductors, an overview of the current problems and different approaches to solve them for the future.

2 Introduction

Gordon Moore's prediction is a law for modern manufacturers of semiconductors [1]. Over many years the development has been oriented on this mentioned law until now. But the current trend looks not so promising to keep this law alive. There are many problems to solve like the power barrier and space problems[2][3][4]. For this reason some organizations were formed to outcome this issue. Two big organizations to mention are:

- ITRS 2.0 (stands for The International Technology Roadmap for Semiconductors), was launched in 2012 and is sponsored by Europe, Japan, USA, Taiwan, Korea
- IEEE RC (stands for The IEEE Rebooting Computing Initiative leverages and IEEE = Institute of Electrical and Electronics Engineers), was also launched in 2012

These scientists are trying to improve performance of PC's and also try to obey Moore's Law. Some good approaches were published by several researchers, which are also listed in this paper.

3 Principle of Semiconductors

SEMICONDUCTORS are electronic components that act like a gate for electrical current. Once there is enough voltage on the component the flow of current is possible and can further be supplied to the rest of a electric circuit for instance. The most basic semiconductor known in the field of electrical engineering is a device called diode. It acts as described previously and can be found in almost every electronic device.

3.1 Structure of a Semiconductor

Every semiconductor consists of at least one p-doped material and one n-doped material. Doping a material means adding an atom of an other kind to a material which only consists of only one type of atoms in order to achieve certain properties. In this case the base form consists of silicon atoms which have four outer electrons which are providing a covalent binding to four other silicon atoms.
To obtain the desired properties an atom with five outer electrons is added to the n-doped material and an atom with three outer electrons to the p-doped material. In this case n-doped means that a negative charged electron is able to move freely because its not strongly bound to another atom. If voltage is applied to this construction then the free electron leaves its position leaving a positive charge behind. The counterpart is p-doped standing for the so called positively charged hole that is considered mobile. In case a voltage is applied as described earlier, an electron from a covalent binding will leave its place and will fill the positive hole, leaving behind another positive hole elsewhere and a negatively charged atom on the current position where the moving electron found its place.

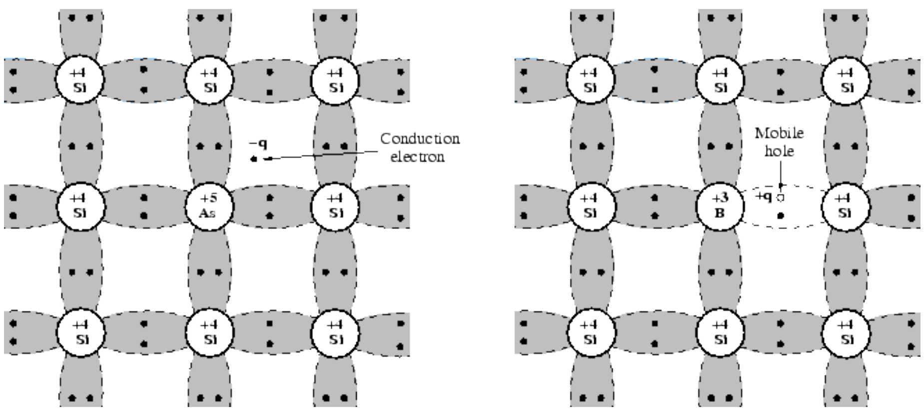

Figure 1: Left P-doped Silicon, Right N-doped Silicon [5]

3.2 The PN-Junction

When these two materials are put together a phenomenon called diffusion causes on one hand electrons on the other hand the positive holes to move to the other material. This leaves a negatively charged p-material and a positively charged n-material behind which causes an electrical field.

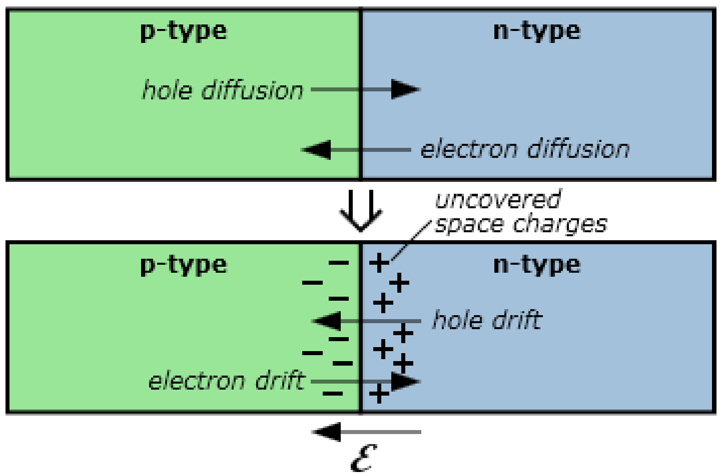

Figure 2: Diffusion and Current Drift

In a state where both materials now hold a charge, positive holes and electrons are compelled towards the opposite charge which causes a movement called current drift. Both the current drift and diffusion cause balance between the electrons and holes which are then not moving anymore. To overcome this lock, energy in form of electrical current is applied to this arrangement to nullify the drift current. If the energy is applied in wrong formation it will amplify the effect of the drift current. In this form adding galvanized contacts to the material would result in a basic semiconductor known as diode.

3.3 The Transistor

An advanced form of a diode is called transistor which can be put into two main categories: a bipolar junction transistor and a field-effect transistor. The main roles of a transistor are to amplify current and to act as a controllable switch.

> *The transistor is a multifunction semiconductor device that, in conjunction with other circuit elements, is capable of current gain, voltage gain, and signal-power gain. The transistor is therefore referred to as an active device whereas the diode is passive.[6]*
>
> *(D. A. Neamen, Page 367)*

In further analysis we will be looking at metal–oxide–semiconductor field-effect (MOS-FET) transistors because they are more advantages in the field of computer-chips.

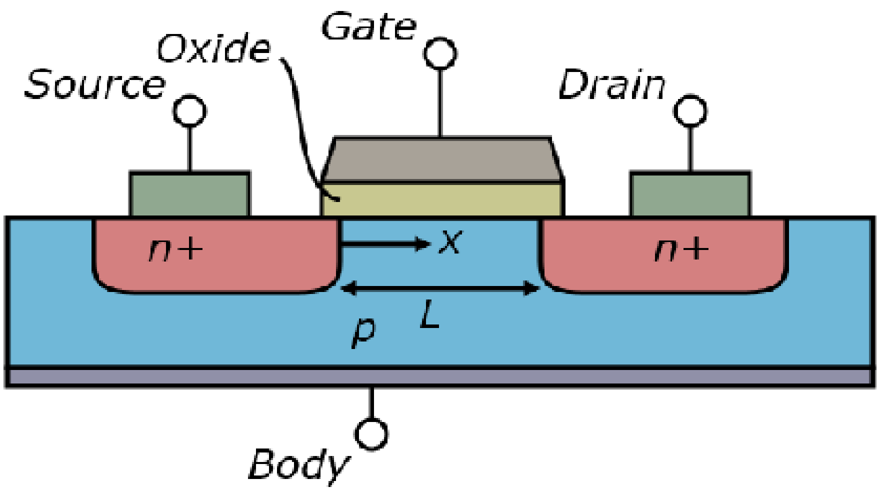

Figure 3: Structure of NPN-Junction of a MOS-FET-Transistor

As seen in Figure 3 it also consists of p and n-doped materials to gain its semiconductor characteristics.

> *The MOSFET, in conjunction with other circuit elements, is capable of voltage gain and signal-power gain. The MOSFET is also used extensively in digital circuit applications where, because of its relatively small size, thousands of devices can be fabricated in a single integrated circuit.[6]*
>
> *(D. A. Neamen, Page 449)*

A massive amount of these arrangements are used in computer-processor chips for switching purposes and without them the advanced computation power on relatively small space that we have nowadays would be unimaginable. In comparison with technology in the past the processor Intel 4004[7] by Intel(1970) had an amount of 2300 transistors on an area of 12 mm^2 where as a modern processor for example the Core i7[8] by Intel(2010) holds 731,000,000 transistors on a space of 263 mm^2.

4 Von-Neumann Architecture

The Von-Neumann-Arichtecture is a computer architecture that was named after the famous mathematician and physicist John von Neumann. Even though it was introduced in 1945 it still provides a foundation for modern day computer-systems.

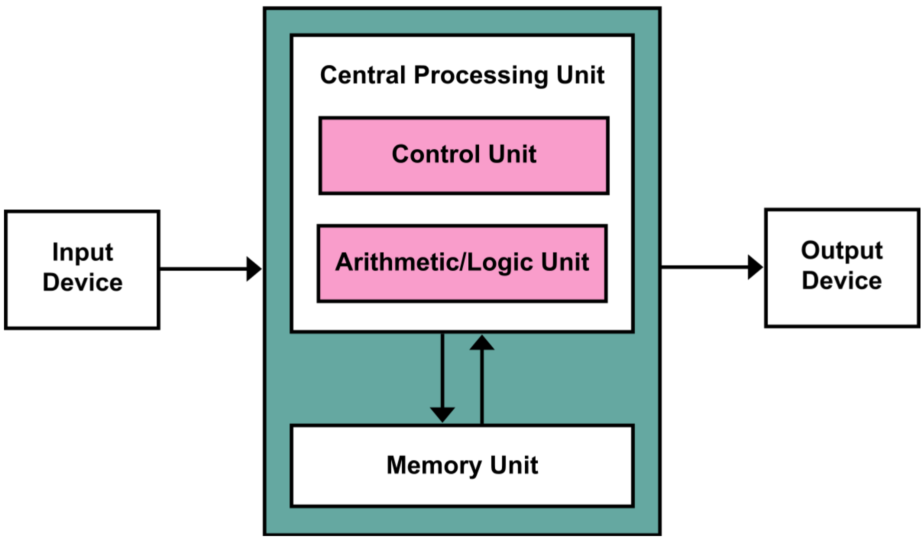

Figure 4: Von-Neumann Architecture 1945 [9]

This model describes a system that has following components:

- Input Device: input form of any kind
- Input Device: output form of any kind
- Central Processing Unit (CPU) which consists of
- Control Unit: to control the flow of instructions
- Arithmetic/Logic Unit: for calculations
- Memory Unit: to store data and instructions

Due to this architecture an instruction fetch and a data operation cannot appear at the same time, because they use the same bus system. This leads to a bottleneck of this system which lead to the Harvard Architecture.

> *Surely there must be a less primitive way of making big changes in the store than by pushing vast numbers of words back and forth through the von Neumann bottleneck. Not only is this tube a literal bottleneck for the data traffic of a problem, but, more importantly, it is an intellectual bottleneck that has kept us tied to word-at-a-time thinking instead of encouraging us to think in terms of the larger conceptual units of the task at hand. Thus programming is basically planning and detailing the enormous traffic of words through the von Neumann bottleneck, and much of that traffic concerns not significant data itself, but where to find it.[10]*
>
> *(John Backus , Pages 613-641)*

5 Moore's Law

In 1965 Gordon Moore, an important personality in the field of computer science, made a very popular prediction called "Moore's Law". Moore as the co-founder and chairman emeritus of Intel Corporation basically stated, that the components on an so called integrated circuit (= IC) will double every year for the following ten years. In 1975 he rethought his prediction and said, that the doubling of the density of components on an IC will happen every two years. Since then, this statement has been very precise and also has become a law for industries and manufacturers as shown in Figure 5.

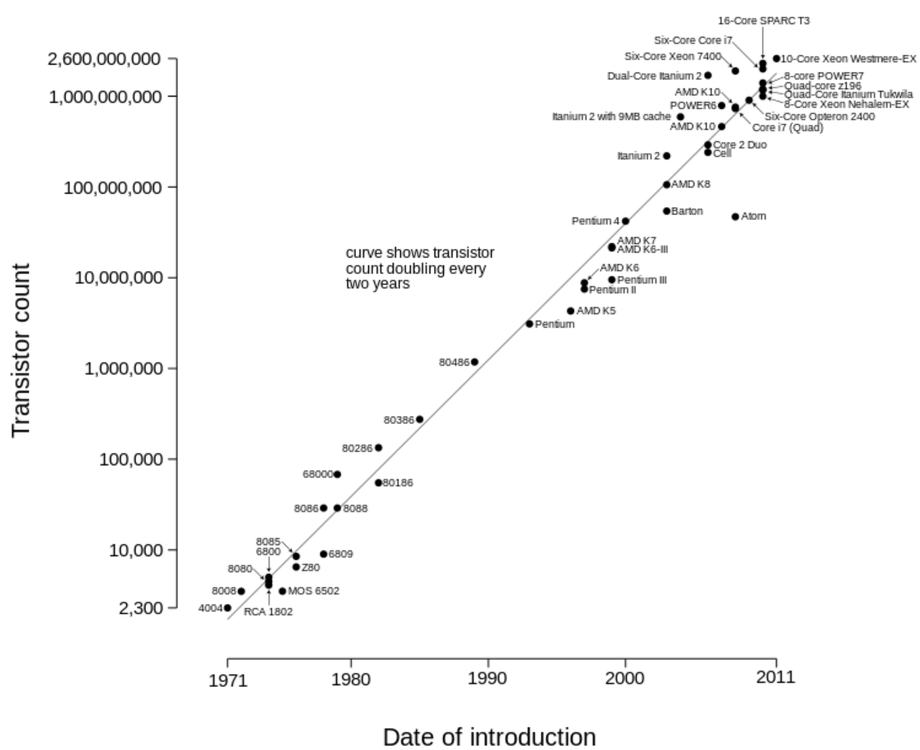

Figure 5: development of the transistor count on a logarithmic scale

6 Problems we are facing

In this section we introduce some problems we are facing at the moment.

6.1 Power limits

In the past the improvement of performance of a computer was easily made by just increasing the clock frequency of the processors. But at some point this wasn't possible anymore, because the community of computer scientists encountered a serious problem. This problem is called "the power barrier" and says, that any operating power that exceeds 100Watts is critical. In this case the processor will just destroy itself, because the cooling techniques at this time were not compatible with the PC hardware. At this point the performance gain by increasing the frequency was no longer an option, as shown in Figure 6.

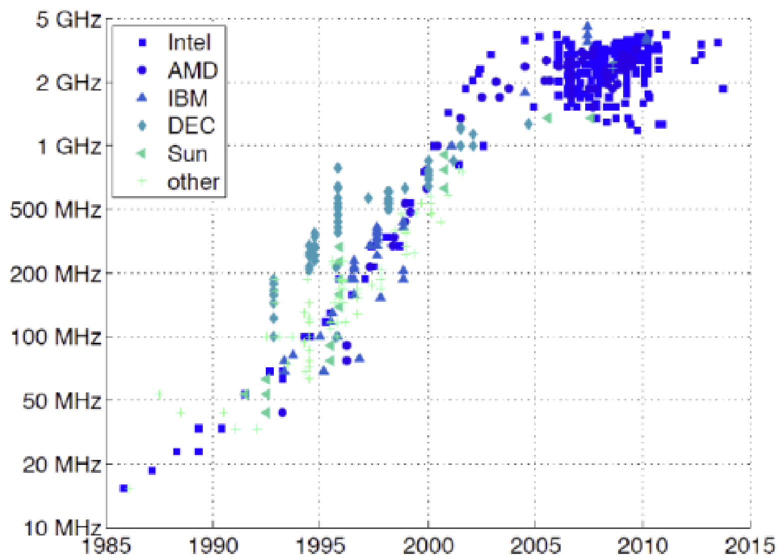

Figure 6: development of frequencies over the years [2]

6.2 Space limits

An other option to improve of performance of a PC is to place more transistors on the same chip, so basically increase the density of transistors on a chip. This worked quite well until now, but now we slowly encounter a problem with this. The semiconductor industry will be approaching transistor measures near to 5nm in the near future, which is the limit of the functionality of a MOS transistor. So also increasing the density will soon be no more a viable option. This problem motivates our next topic: 3D Power Scaling.

7 3D Power Scaling

3D power scaling is a really promising method for the future (some scientists experiment on it right now) to improve PC computations. In this method the planar transistor is rotated by 90 degrees. So we can stack multiple planes of verticals transistors on top of each other, as shown in Figure 7.

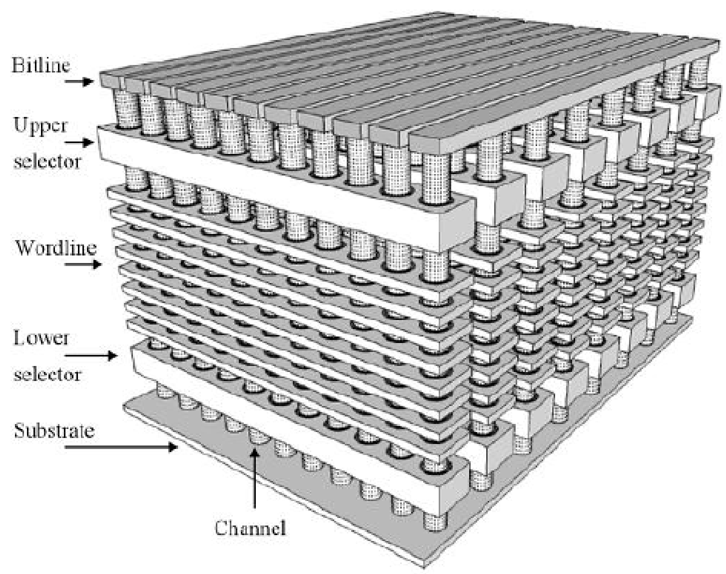

Figure 7: a basic construction of multi planes of transistors [2]

So one can state, that the numbers of transistors will continue to increase, even more than predicted in "Moore's Law".

8 A Look Into The Future

The manufacturers will still improve performance with methods like 3D power scaling, but the issue with ”the power barrier” still remains. In this section we like to shortly introduce some new concepts for computation, which will be probably pioneering for the future of computers.

8.1 Decreasing The Operating Voltage

Let f be the clock frequency of a processor, V the operating voltage and C the capacitance, then the Power P can be calculated by the equation:

$$P = C \cdot V^2 \cdot f$$

Thus lowering the voltage is a valid option for improving the performance.

8.2 Asynchronous circuits

This technique basically describes, that no clock is used to synchronize the elements on a circuit, so they can compute independent from each other. This method has the advantage, that it has a known potential for speedup. But there are several things that need to be explored in this field, like complexity.

8.3 Memory Driven Computation

Memory driven computation is probably one of the most promising approaches, because the data amount that is transported is exploding and there is no end in sight[11]. It's basic concept is to move the calculation away from the CPU to the data itself, as shown in Figure 8.

This method has several advantages:

- very fast
- less complex
- less space needed
- less energy needed
- reduced cost

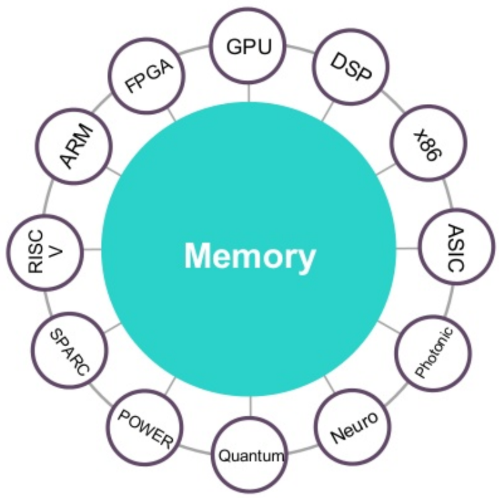

Figure 8: the basic concept of memory driven computation[11]

A rough description of this process is, that we have to combine memory and storage into one environment to improve processing speed and energy efficiency. Distances need to be eliminated by using photonics (optical fiber,...).

9 Conclusion

We are reaching several limits as described in the section "Problems we are facing", so we have to rethink our computation process from scratch. A new and more efficient architecture is needed for the future as well as a new component, which is replacing the current MOS transistor to allow higher operating frequencies. There are good different approaches mentioned in the section "A Look Into The Future" and there are even more. So one can say, that we have a promising future on semiconductors.

References

[1] R. Schaller, "Moore's law: past, present and future," 1997.

[2] C. T.M. and G. P.A., "On the foundation of the new computing industry beyond 2020 executive summary ieee rebooting computing and itrs september 2015," 2015.

[3] S. Greengard, "The future of semiconductors," 2015.

[4] H. Cheong, "Management of technology strategies required for major semiconductor manufacturer to survive in future market, graduate school of management of technology, hoseo university, asan 336- 795, korea, procedia computer science 91," 2016.

[5] R. Wittmann, "Miniaturization problems in cmos technology: Investigation of doping profiles and reliability."

[6] D. A. Neamen, "Semiconductor physics and devices-basic principles-third edition," 2003.

[7] Intel, "4004 single chip 4-bit p-channel microprocessor."

[8] Intel, "Intel® core^{TM} i7-900 desktop processor extreme edition series and intel® core^{TM} i7-900 desktop processor series datasheet, volume 1," 2010.

[9] W. H. Encyclopedia, "Von neumann architecture."

[10] J. Backus, "Can programming be liberated from the von neumann style?: a functional style and its algebra of programs," 1978.

[11] K. Keeton, "Memory-driven computing," 2017.