Lecture 1

Gidon Rosalki

2025-03-23

1 Introduction

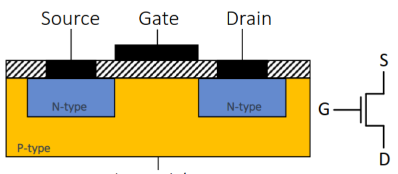
ENIAC - the first computer. Used vacuum tubes, which are gates that operate based off electronic emissions of heated wires. Was about 27 tonnes, 160KW of power, used 18,000 vacuums tubes which needed to be changed at a rate of about 2,000 a day. It was the first computer to find the first 2,000 digits of pi. ”There is no reason anyone would want a computer in their home” - Ken Olsen
Computers consist of IO devices, storage devices which may be volatile (do not survive power disconnection eg RAM), and non volatile (storage, HDDs, SSDs, etc.), the processing unit CPU, the operating system, device drivers, and application software. We will cover the CPU, and application SW, and mention storage devices, and the OS.
Moore’s Law (1965): Intel co-founder Gordon Moore predicted in 1965 that the number of transistors on a chip will double about every 2 years, which has been the case for 40 years now. This was achieved by telling the engineers that they may go home once they have achieved a doubling in 2 years. If they achieved more, then it was spread out to a doubling every 2 years.
The AAA paradigm: Every computer scientist should master Architecture, Algorithms, and Applications.

2 Semiconductors

The first transistors were made in 1947 by John Bardeen and Walter Brattain in Bell Labs, and was made of germanium. A transistor is a solid state switch, which may be switched through the application of electric charge. In 1958 Jack S Kilby at Texas Instruments developed the first integrated circuit (again using germanium). It contained 5 components, of 3 different types: Transistors, resistors, and capacitors.
The periodic table was created first by Mendeleev in 1869. It was a way to organise the chemical elements by weight, but these days is organised by electron configuration. Each row of the table is a ’layer’ of electrons, and for the most part the elements in the same column have the same number of electrons in their outer shell. Bohr suggested the Bohr model, where the nucleus contains the protons and neutrons, with electrons in shells around it, where the first shell has up to 2, second and third up to 8, and the fourth and fifth up to 18, and so on.
Atoms ”want” a full outer shell, and this is achieved through forming bonds. These can be covalent bonds, where an electron is shared between the atoms, and ionic bonds, which bonds charged ions together. (There is also metallic bonding not discussed in this course).
We consider that electrical current is the flow of electrons (or holes). The electrons are viewed by convention as moving from positive to negative. This is analogous to water flow. Voltage is the potential difference. The units are volts, and we use the term ”ground” to indicate the point of lowest potential. This is also analogous to water potential energy (consider a water tower instead of a battery). The charge is the amount of electrons that are stored, and capacitance is the ability to store charge. When a charge is stored in a capacitor, a voltage is applied on it. A capacitor is analogous to a water tower. In an electrical circuit current will flow from high potential to lower potential. Resistance will impact the rate of flow, higher resistance slows the flow, where lower resistance allows for greater flow. Electrical circuits may also include branches. Flow is either from a power supply, of from a capacitor to ground.
We will use capacitors in order to hold 0 or 1. A 1 is represented by a capacitor with high voltage, and 0 is low voltage. Logical values will be represented with 0s and 1s in binary, and we will generally use hexadecimal to represent groups in a short format.
We use silicon to create our chips. Pure silicon is a semiconductor and does not conduct electrical current well because it has few free charge carriers. Silicon dioxide (SiO_2, or sand) is an insulator, and conducts no electrical current. We use doping, an intentional introduction of impurities into a semiconductor to modulate its properties. We create P-type, and N-type. P-type has extra aluminium in the semiconductor, which reduced the number of available electrons, meaning that there are extra holes. N-type has added phosphorous, which is an electron donor, and increases the number of free electron in the substance.
If we connect together p-type and n-type, we can create a PN junction, which conducts in one direction, and blocks current in the other direction, which we call a **diode**. Electron flow will occur from the n-type to the p-type, meaning conventional current will flow from the p-type to the n-type.

3 MOS-FET transistor structure and operation

MOS-FET = Metal Oxide Semiconductor - Field Effect Transistor.



N-Channel / NMOS

Contains the terminals Source, Drain, and Gate. They are constructed from an insulator, and two NP junctions, and a capacitor which is the gate body. There are two variates, N-Channel (NMOS), and P-Channel (PMOS).
When considering transistors, there are a few parts to consider. L is the length of the channel / gap between the source and gate. W is the width of the transistor (along the source/gate), and Lox is the oxide thinness, which impacts the gate capacitance.
NMOS is used as a voltage controlled switch. The gate controls whether current can flow between the source and the drain. When the gate reads 0, we have two charged NP junctions, and thus current cannot flow between S and D. When the gate reads 1, then there is an N like channel between S and D, and the current can flow. PMOS operates in the opposite direction.

4 Manufacture

Making wafers is done by melting pure silicon, and adding a very small amount of the dopant. You then insert and rotate an ingot, a starting crystal on which the doped silicon crystallises. Once this process is completed, the crystals are cut and polished into wafers. To process the wafers, geometries are defined using light, and a film coating that is light sensitive. Generally uses a wavelength of $< 1nm$. Each step creates a mask, does a chemical/physical step on the light area (etching, doping, depositing), clean the mask, and then flattened.