COMPUTER ORGANIZATION AND ARCHITECTURE

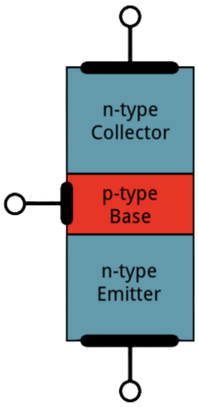
"In 1969 the world recorded the first man to step on the moon, Neil Armstrong with his co-astronauts. The National Aeronautics and Space Administration (NASA) used the Apollo Guidance Computer with RAM of 4KB, a 32KB hard disk. Geometrically they measured 60cm x 30cm x 15cm and physically weighed around 30kg. Surprisingly we now store an INTEGER (int) in a computer program!"

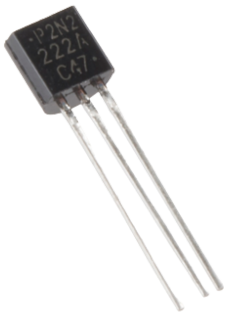
OVERVIEW

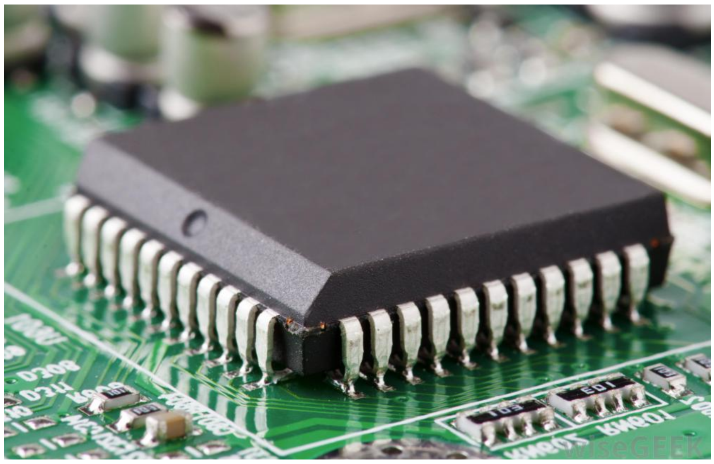
In order to understand the rationale behind such astronomical change one has to explore the root level cause, The TRANSISTORS and INTEGRATED CIRCUITS and their evolution.

TRANSISTOR
"A transistor is a three-terminal, bipolar, current-controlled semiconductor device with three parts: Emitter, Base, and Collector". The emitter zone is strongly doped, whereas the collector region has a large volume with medium doping levels.
Transistor was invented in the late 1947s, with Silicon Atom as the primary production component.

INTEGRATED CIRCUITS
In 1958, a monolithic integrated circuit (IC) was created by assembling a large number of transistors on a single chip. Modern Digital Electronics has been transformed by the introduction of the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) and the rapid development of ICs that are geometrically smaller, quicker, and less expensive.





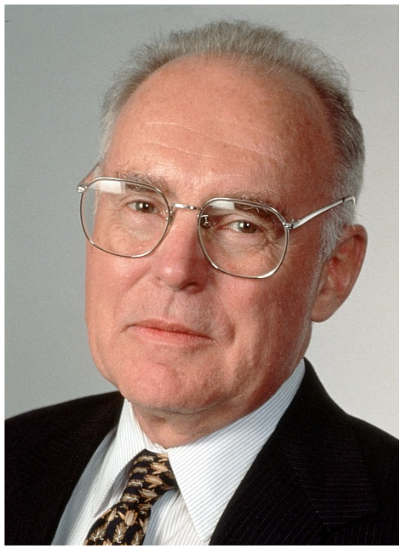


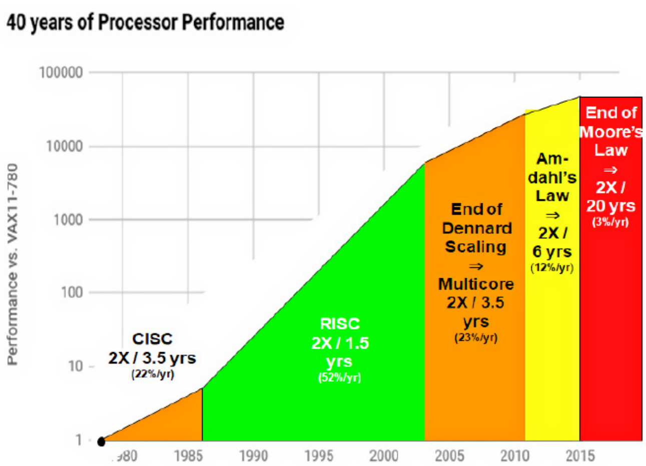
MOORE'S LAW

Moore's law was initially published in Electronics magazine in 1965, when Moore was a founding member of Fairchild Semiconductor and head of research. Moore's law states that the number of components on a computer chip doubles every two years. Moore anticipated that every two years, the number of transistors on a single square inch of an integrated circuit chip would double.

Moore's law is closely related to MOSFET scaling, as the rapid scaling and miniaturization of MOSFETs is the key driving force behind Moore's law. Mathematically, Moore's Law predicted that transistor count would double every 2 years due to shrinking transistor dimensions and other improvements.

"The complexity for minimum component costs has increased at a rate of roughly a factor of two per year. Certainly over the short term this rate can be expected to continue, if not to increase. Over the longer term, the rate of increase is a bit more uncertain, although there is no reason to believe it will not remain nearly constant for at least 10 years."

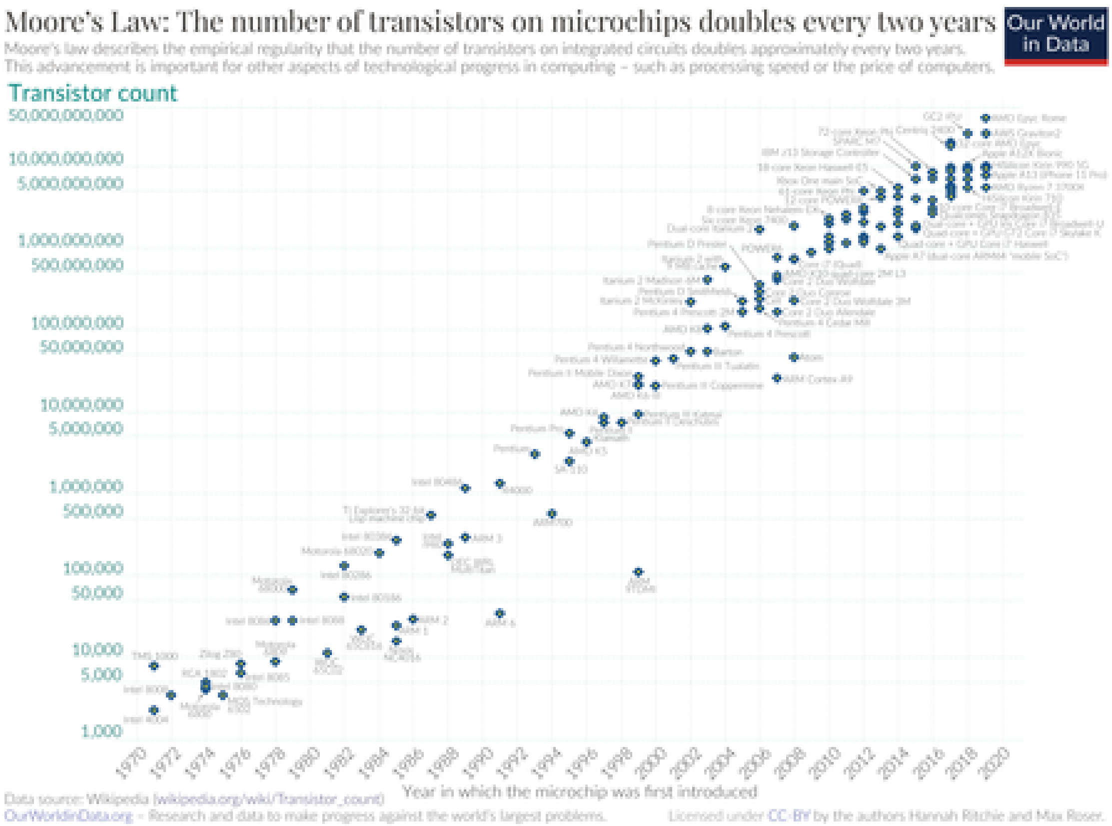




SURVEY

Since the prediction made in 1965 to till date, the world Electronics community has almost obeyed the law.
The following data depicts the temporal distribution of size of MOSFETs.

10 µm - 1971
6 µm - 1974
3 µm - 1977
1.5 µm - 1981
1 µm - 1984
800 nm - 1987
600 nm - 1990
350 nm - 1993
250 nm - 1996
180 nm - 1999
130 nm - 2001
90 nm - 2003
65 nm - 2005
45 nm - 2007
32 nm - 2009
22 nm - 2012
14 nm - 2014
10 nm - 2016
7 nm - 2018
5nm - 2020



- The significance of this rule is underscored by the fact that it has resulted in a technical shift from microelectronics to nanoelectronics, as well as the creation of a new industrial segment called "nanotechnology," which is growing at an exponential rate. This movement has sparked a surge of interest in new sectors such as nanomaterials and new semiconductor manufacturing optimization methods.
- One of the law's economic effects is that computer devices continue to expand in complexity and processing power at an exponential rate while lowering costs for both the producer and the consumer.

SCALING

The reduction of all Geometric dimensions of a chip by a factor(s) where s is an integer is called Scaling.

Geometric proportions:
Geometrical (constant field) scaling is the process of decreasing the horizontal and vertical physical feature sizes of on-chip logic and memory storage functions in order to increase density (lower cost per function) and performance (speed, power).

Scaling to Equivalents:
Equivalent scaling (which happens in tandem with and permits continuing geometrical scaling) refers to advancements in 3-dimensional device structure ('Design Factor') as well as other non-geometrical manufacturing methods and novel materials that impact the chip's electrical performance.

Equivalent Scaling in Design:
Design Equivalent Scaling refers to design technologies that provide high performance, low power, high reliability, cheap cost, and high design productivity. It happens in conjunction with equivalent scaling and continuing geometric scaling.

SEMICONDUCTOR LITHOGRAPHY

Semiconductor Lithography is another crucial factor that will lead to large-scale MOSFETS integration.

Semiconductor Lithography, also known as Photolithography, literally translates to "Light-Stone-Write." Photolithography is a microfabrication technique that involves transferring geometric patterns to a film or substrate.
Over the last three decades, lithography advancements and device scaling have been the key drivers of industrial productivity, as anticipated by Moore's law.

Following table gives a comprehensive impact on Transistor multiplicity and adhering to Moore's law due to Scaling and Photolithography

	Introduction Date	Clock Speeds	Bus Width	Number of Transistors	Addressable Memory	Virtual Memory	Brief Description
4004	11/15/71	108 KHz	4 bits	2,300 (10 microns)	640 bytes		First microcomputer chip, Arithmetic manipulation
8008	4/1/72	108 KHz	8 bits	3,500	16 KBytes		Data/character manipulation
8080	4/1/74	2 MHz	8 bits	6,000 (6 microns)	64 KBytes		10X the performance of the 8008
8086	6/8/78	5 MHz 8 MHz 10 MHz	16 bits	29,000 (3 microns)	1 Megabyte		10X the performance of the 8080
8088	6/1/79	5 MHz 8 MHz	8 bits	29,000 (3 microns)			Identical to 8086 except for its 8-bit external bus
80286	2/1/82	8 MHz 10 MHz 12 MHz	16 bits	134,000 (1.5 microns)	16 Megabytes	1 gigabyte	3-6X the performance of the 8086
Intel386(TM)DX Microprocessor	10/17/85	16 MHz 20 MHz 25 MHz 33 MHz	32 bits	275,000 (1 micron)	4 gigabytes	64 terabytes	First X86 chip to handle 32-bit data sets
Intel386(TM)SX Microprocessor	6/16/88	16 MHz 20 MHz	16 bits	275,000 (1 micron)	4 gigabytes	64 terabytes	16-bit address bus enabled low-cost 32-bit processing
Intel486(TM)DX Microprocessor	4/10/89	25 MHz 33 MHz 50 MHz	32 bits	1,200,000 (1 micron, .8 micron with 50 MHz)	4 gigabytes	64 terabytes	Level 1 cache on chip
Intel486(TM)SX Microprocessor	4/22/91	16 MHz 20 MHz 25 MHz 33 MHz	32 bits	1,185,000 (.8 micron)	4 gigabytes	64 terabytes	identical in design to Intel486(TM) DX but without math coprocessor
Pentium® Processor	3/22/93	60MHz 66MHz 75MHz 90MHz 100MHz 120MHz 133MHz 150MHz 166MHz	32 bits	3.1 million (.8 micron)	4 gigabytes	64 terabytes	superscaler architecture brought 5X the performance of the 33-MHz Intel486 DX processor

The above data is an extract from the publishing "**Moore's law governs the silicon revolution**" by Probir K. Bondyopadhyay in 1996.
Following data includes recent developments and trends

2019: 20 billion Transistors in Samsung V-Nand Chip.

2020: 26 billion Transistors in Water Scale Engine 2.

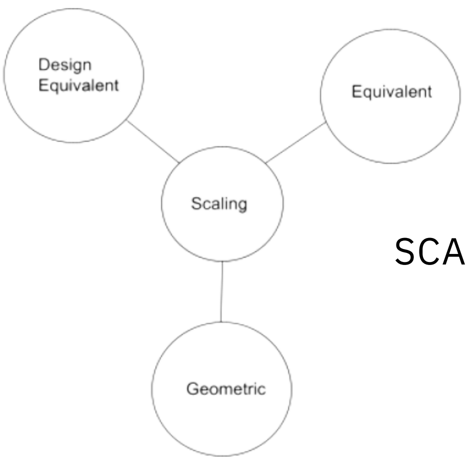
2021: 59 billion MOSFET AMD's instinct.

2022: 114 billion Apple's ARM based MI Ultra.

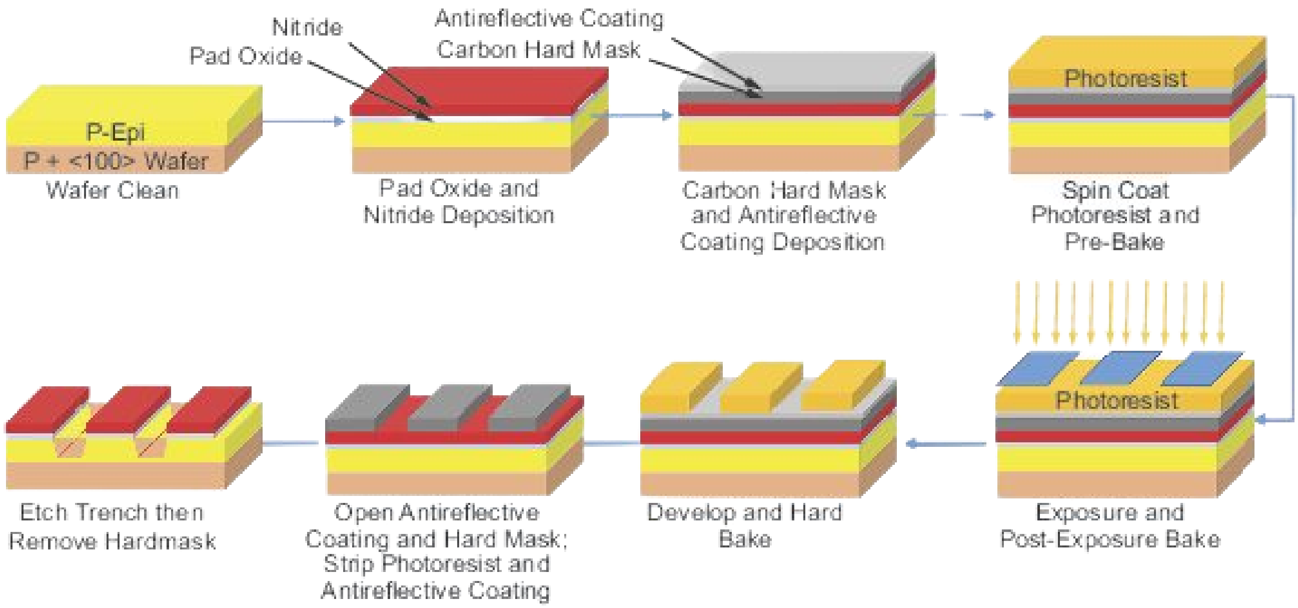
- A team of Korean researchers from the Korea Advanced Institute of Science and Technology (KAIST) and the National NanoFab Center produced a 3 nm transistor based on FinFET technology in 2006, making it the world's smallest nanoelectronics device at the time.
- In 2012, a team of researchers from the University of New South Wales reported the creation of the world's first functional transistor, which consisted of a single atom precisely positioned in a silicon crystal (not just picked from a large sample of random transistors). Moore's law anticipated that ICs in the lab will achieve this milestone by 2020.
- The precision and correctness of Moore's forecast, which has been turned into a law, is depicted by the improvisation and sophisticated integration of massive technology into nanochips.
- Improvisation and complex integration of gigantic technologies onto nanochips portrays the precision and the accuracy of Moore’s prediction turning it into a Law.

RECENT TRENDS

- Microprocessor architects report that since around 2010, semiconductor advancement has slowed industry-wide below the pace predicted by Moore's law. The rate of improvement in physical dimensions known as Dennard scaling also ended in the mid-2000s. As a result, much of the semiconductor industry has shifted its focus to the needs of major computing applications rather than semiconductor scaling.
- The explosion of hyperconnectivity, big data, and artificial intelligence applications has increased the pace of innovation and the need for “Moore’s law-style” improvements in delivered technology.



SCALING



SEMICONDUCTOR LITHOGRAPHY

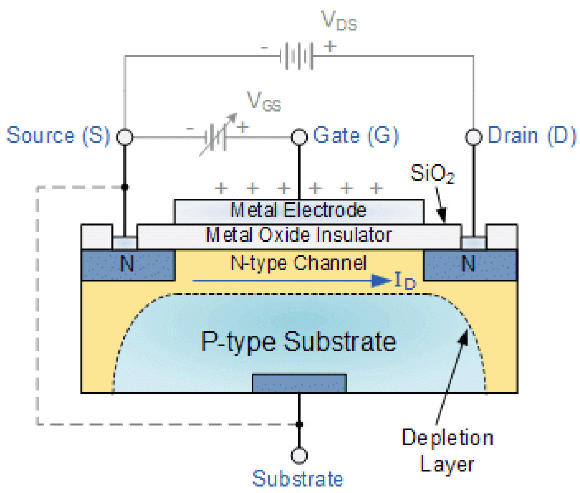
HOW DID MOORE LAW TURNED TO REALITY?

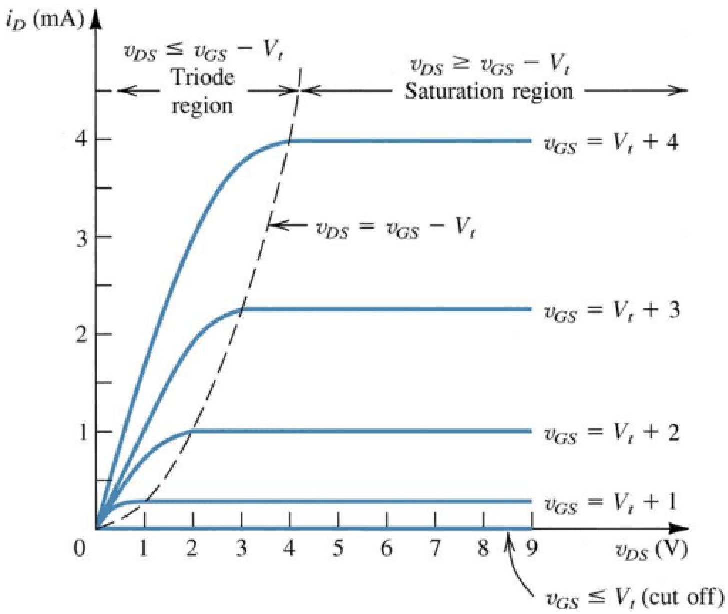
- Moore's law isn't really a law in the legal sense or even a proven theory in the scientific sense. Many new types of transistors were designed after 1970 in order to improve efficiency, close existing loopholes, and manufacture more transistors onto the chip for better CPUs.
- The developed versions with various parameters for upgrading bipolar junction transistors are listed below.

MOSFET

- A Metal Oxide Semiconductor Field-Effect Transistor is a four-terminal device with four terminals: source (S), gate (G), drain (D), and body (B). The body of the MOSFET is connected to the source terminal, resulting in a three-terminal device similar to a field-effect transistor.
- The electrical fluctuations in the channel width, as well as the flow of carriers, determine the operation of a MOSFET (either holes or electrons). Charge carriers enter the channel through the source terminal and escape through the drain terminal.

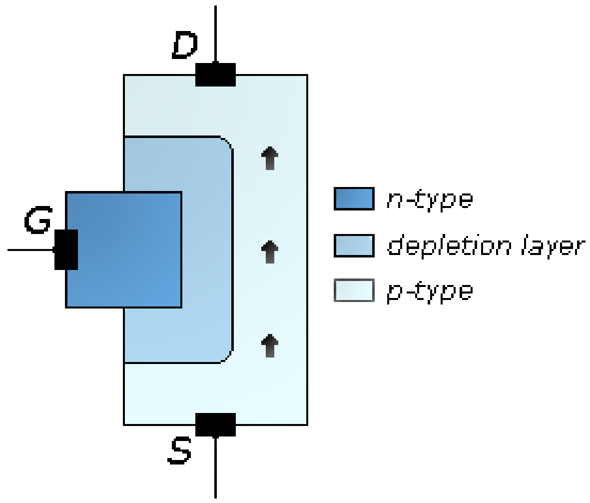
- The voltage on an electrode called the gate, which is placed between the source and the drain, controls the width of the channel. An exceedingly thin coating of metal oxide separates it from the channel.
- The device's MOS capacity is the critical region where the complete operation is carried out.

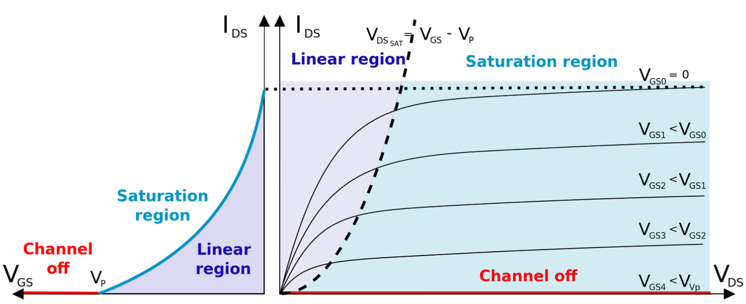




JFET

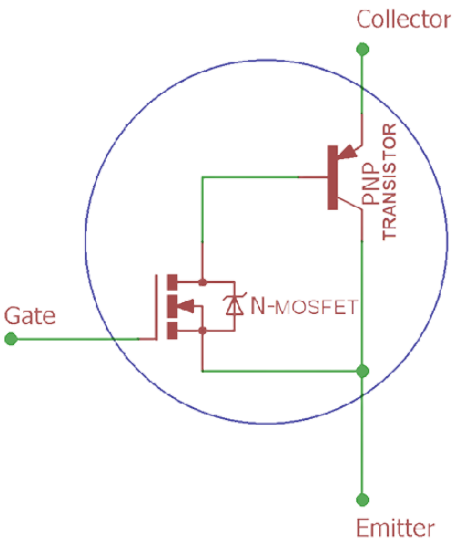
- One of the most basic varieties of field-effect transistor is the junction-gate field-effect transistor (JFET). JFETs are three-terminal semiconductor devices that can be used to produce amplifiers or as electronically controlled switches or resistors.
- JFETs, unlike bipolar junction transistors, are only voltage-controlled and do not require a biasing current.



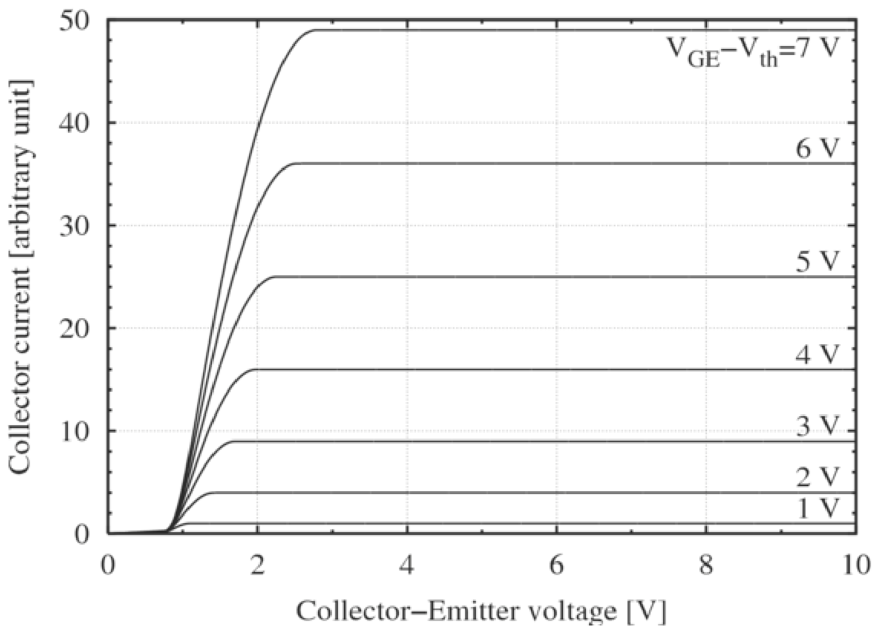


IGBT

- An insulated-gate bipolar transistor (IGBT) is a three-terminal power semiconductor device principally used as an electronic switch that has evolved to combine high efficiency and quick switching as it has been developed. A metal–oxide–semiconductor (MOS) gate structure controls four alternating layers (P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P–N–P.
- The advent of the IGBT has accelerated the Power electronics industry, which has considerably contributed to huge Electronic and Electrical systems. Electronic BGTs functioned between 0V-20V.

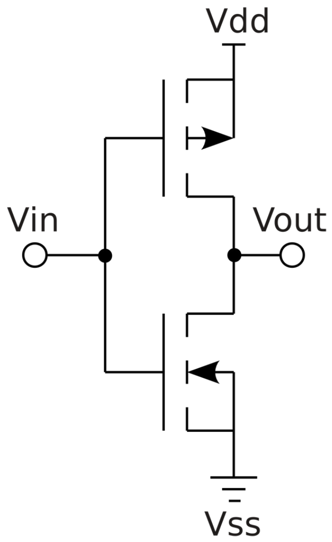


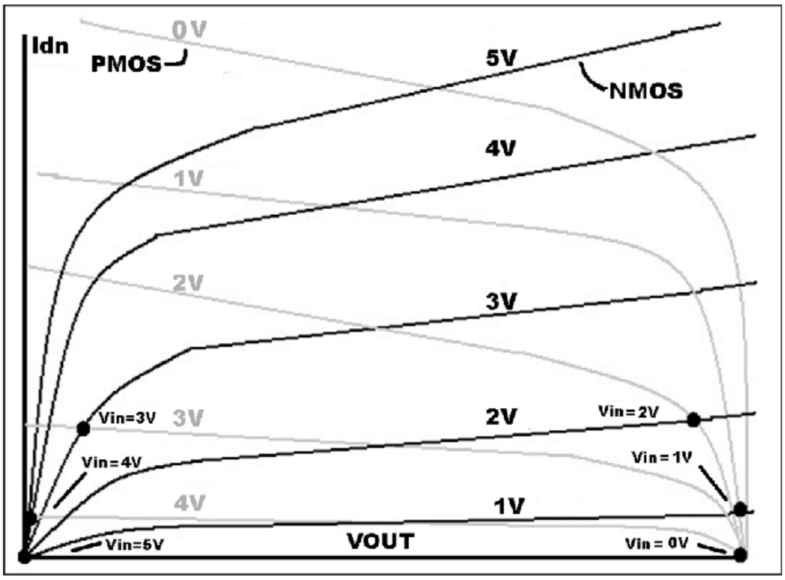
Equivalent Circuit



CMOS

- Static power dissipation is nearly non-existent in a complementary MOS circuit. Only if the circuit really switches is power squandered. This enables the integration of more CMOS gates on an IC than is possible with NMOS or bipolar technology, resulting in much improved performance.
- P-channel MOS (PMOS) and N-channel MOS (NMOS) are complementary Metal Oxide Semiconductor transistors (NMOS).

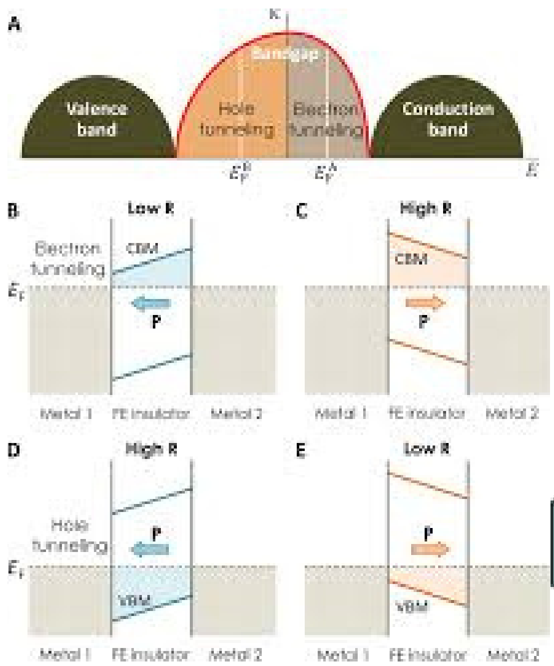


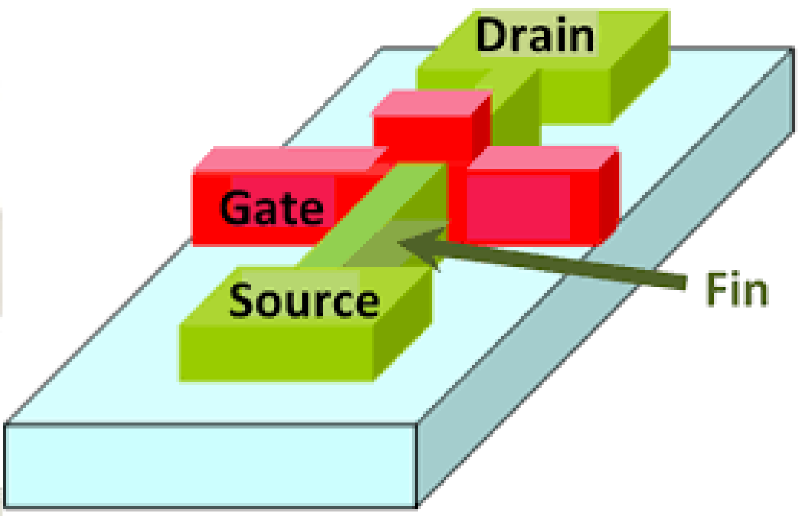


ILL EFFECTS OF MOORE LAW

QUANTUM TUNNELING

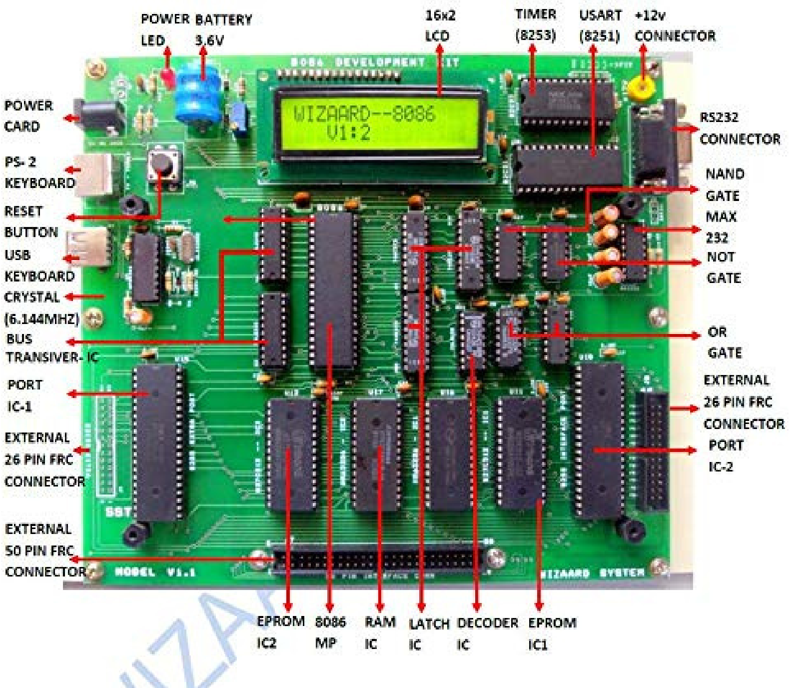
- The distance between the transistors' source and drain is so short that electrons jump over the barrier. As a result, rather of remaining in the desired logic gate, electrons flow continually from one gate to the next, thereby rendering the transistor inoperable.
- A dramatic redesign of the transistor, i.e. The FinFET, is the most popular nanoscale transistor, is required to overcome this. On three sides of the channel, the FinFET has gate dielectric. The gate-all-around MOSFET (GAAFET) structure, in comparison, offers even greater gate control.





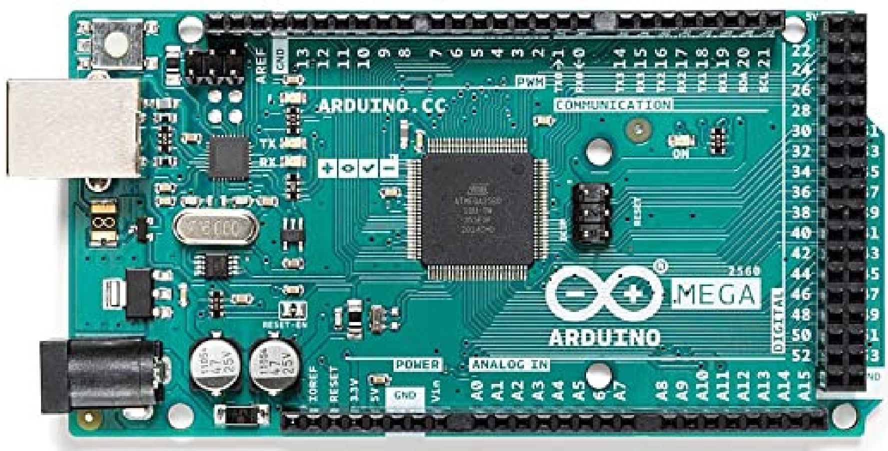
A SIMPLE COMPARISON

MICROCONTROLLER VS ARDIUNO



8086 MICROCONTROLLER

ARDIUNO



SIMULATING SUM OF TWO NUMBERS

MASM CODE FOR 8086 MICROCONTROLLER:

```
data segment
a db 09h
b db 02h
c dw ?
data ends

code segment
assume cs:code,ds:data
start:
mov ax,data
mov ds,ax
mov al,a
mov bl,b
add al,bl
mov c,ax
int 3
code ends
end start
```

EMBEDDED C CODE FOR ARDIUNO:

```
int a = 2;
int b = 7;
int sum;
sum = a + b;
```

INTERPRETATION

PARAMETER	8086 MICROCONTROLLER	ARDIUNO
SIZE	29,000 MOS Transistor Count	TRANSISTOR COUNT = 1,00,000
CLOCK SPEED	5MHz	48MHz
EFFICIENCY	LESS EFFECIENT	MORE EFFECIENT

EXTENDING MOORE LAW

Unfortunately, the rate of technological advancement has not decreased, and we are approaching the conclusion of Moore's Law. Many in the business are hesitant to make firm predictions about when the concept will become obsolete, but the general assumption is that it will happen between 2020 and 2025 — and no one appears to have a viable answer to this technological stumbling block.

AI CHIPS

- Artificial intelligence chips (also known as AI hardware or AI accelerators) are specially built accelerators for ANN-based applications.
- Switching to a more ASIC-focused design might provide a stopgap solution and extend Moore's Law by about 10 years, but it would increase the von Neumann bottleneck in the industry. Modern circuits are capable of operating so quickly that the time it takes for information to move between chips wastes computation time.

3D CHIPS

- Three-dimensional integrated circuits, in which many wafers are stacked on top of one other and linked vertically. Design engineers can lower the distance data has to travel while simultaneously reducing chip footprint by successfully integrating numerous ICs together. 3D integration covers a wide range of technologies, including 3D wafer-level packaging (3DWLP), 2.5D and 3D interposer-based integration, 3D stacked ICs (3D-SICs), monolithic 3D ICs, 3D heterogeneous integration, and 3D systems integration.
- This might extend the usefulness of existing computer technologies, allowing Moore's Law to continue to apply.

CONCLUSION

When we look back at NASA's Apollo 11 mission and Neil Armstrong's success tale, we can see a tremendous growth in the quantity of transistors and a drop in their size. Again, Moore's law was only a forecast in 1965, and it is clear that the subsequent trend has followed it.
As we traverse and climb the Computer Science incline, it falls to us to remain abreast of current changes in the global electronics order. The World Electronics and Computer Science community will be affected by socio-political, socio-economic, and techno-economic developments in addition to science and technology.

Solutions to generic issue statements are headed towards the period of Industrial Revolution 4.0, often known as the Digital Revolution.

Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Internet Of Things (IoT), and other terms are used to describe these technologies. Such solutions frequently need advanced computational abilities and substantial power usage.
Electronics such as Very-Large-Scale-Integration (VLASI), Very-Very-Large-Scale-Integration (VVLSI), and others are enabling more computing.

REFERENCES

1. Bondyopadhyay, P.K., 1998. Moore's law governs the silicon revolution. Proceedings of the IEEE, 86(1), pp.78-81)

2.Brown, K.H., 2000, July. Moore's law-is there more?. In Digest of Papers Microprocesses and Nanotechnology 2000. 2000 International Microprocesses and Nanotechnology Conference (IEEE Cat. No. 00EX387) (pp. 2-3). IEEE)

3. Kahng, A.B., 2010. Scaling: More than Moore's law. IEEE Design & Test of Computers, 27(3), pp.86-87)

4. Mulay, A., 2015. Sustaining Moore's Law: Uncertainty Leading to a Certainty of IoT Revolution. Synthesis Lectures on Emerging Engineering Technologies, 1(1), pp.1-109.

5. Adee, S., 2008. The data: 37 Years of Moore's Law. Ieee Spectrum, 45(5), pp.56-56.