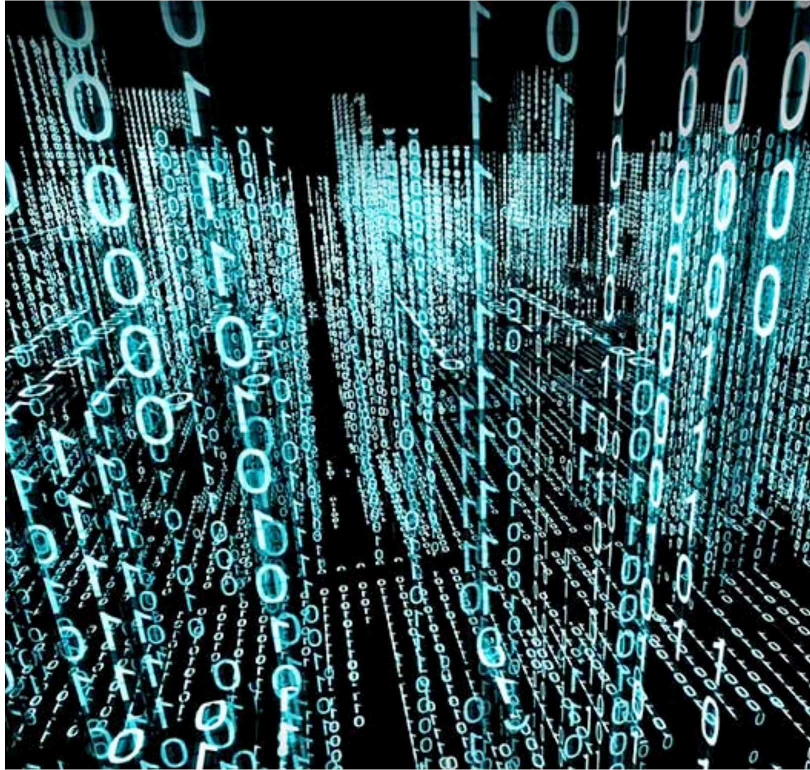


EECS 151/251A

Spring 2024

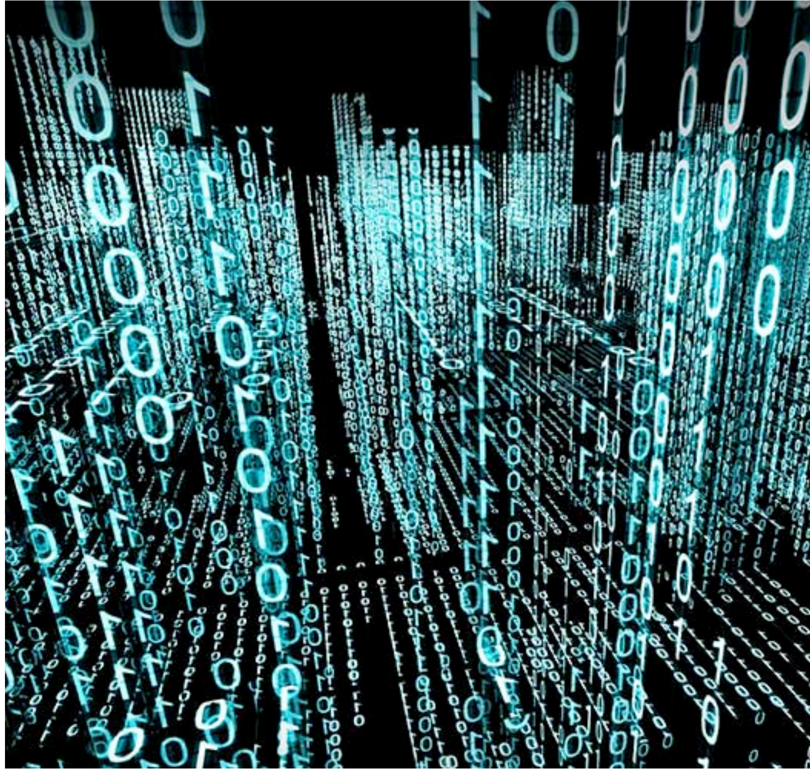
Digital Design and Integrated Circuits

Instructor: Wawrzynek

Lecture 9: CMOS1

Announcements

- Monday is an Academic Holiday
- New Midterm Date/Time:
  - **Tue Mar 12 2024 7:00-10:00PM**
  - **MOFF101, VLSB2040**
- Lecture Schedule change:
  - Monday Mar 11, in-class MT review



From the Bottom Up

IC processing

MOS transistors

CMOS Circuits

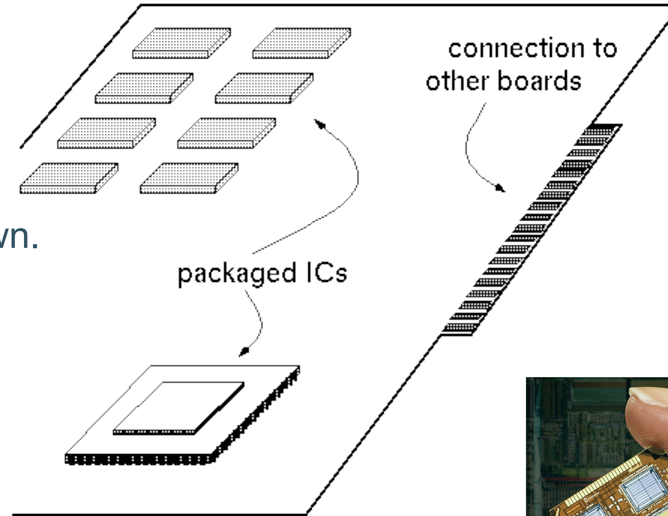
Overview of Physical Implementations

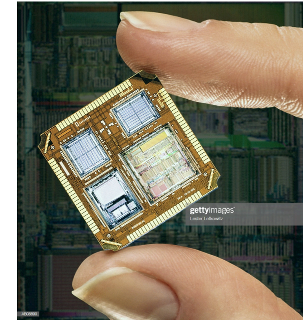
The stuff out of which we make systems.

- Integrated Circuits (ICs)
  - Combinational logic circuits, memory elements, analog interfaces.
- Printed Circuits (PC) boards
  - substrate for ICs and interconnection, distribution of CLK, Vdd, and GND signals, heat dissipation.
- Power Supplies
  - Converts line AC voltage or battery DC voltage to regulated DC low voltage levels.
- Chassis (rack, case, ...)
  - holds boards, power supply, fans, provides physical interface to user or other systems.
- Connectors and Cables.
- Peripheral and I/O components.

Printed Circuit Boards

- fiberglass or ceramic
- 1-25 conductive layers
- ~1-20in on a side
- IC packages are soldered down.

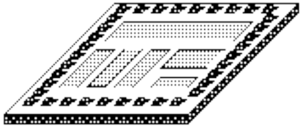




Multichip Modules (MCMs) / “chiplets”

- *Multiple chips directly connected to a substrate. (silicon, ceramic, plastic, fiberglass) with or without chip packages.*

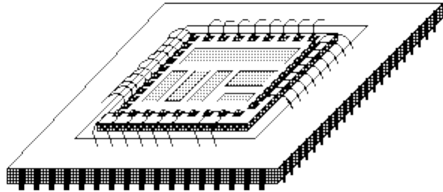
Integrated Circuits



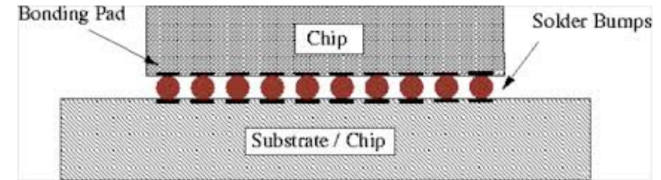
- Primarily Crystalline Silicon
- 1mm - 25mm on a side
- 100 - 20B transistors
- (25 - 250M "logic gates")
- 3 - 10 conductive layers
- state-of-the-art feature size 5nm = 0.005×10^{-6} m
- "CMOS" most common - complementary metal oxide semiconductor

<u>*Chip in Package*</u>

- *Package provides:*
  - *spreading of chip-level signal paths to board-level*
  - *heat dissipation.*
- *Ceramic or plastic with gold wires.*



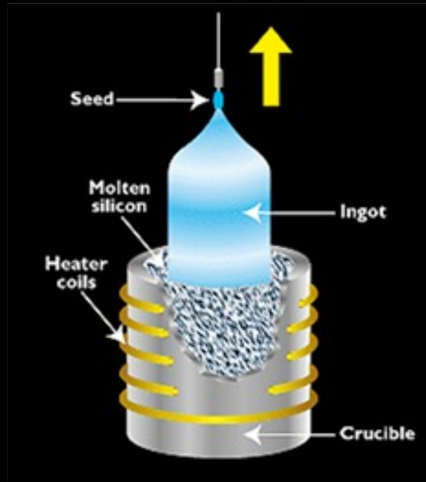
wire bond



bump bond

Chip Fabrication

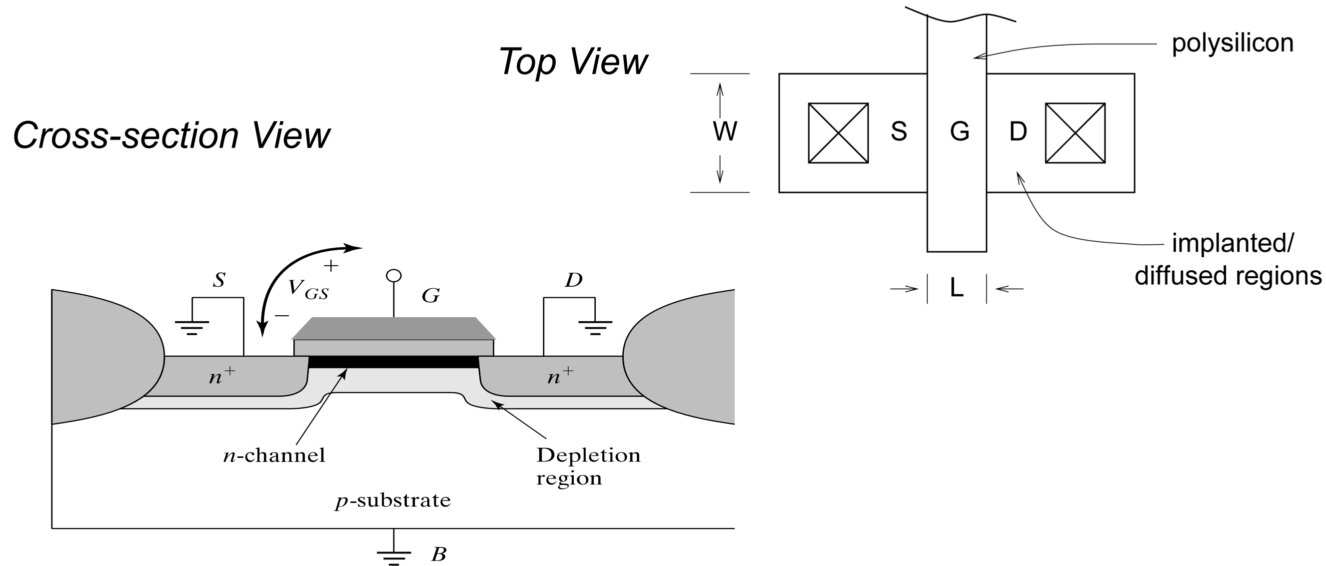
Silicon "ingots" are grown from a "perfect" crystal seed in a melt, and then purified to "nine nines".



Ingots sliced into 450μm thick wafers, using a diamond saw.

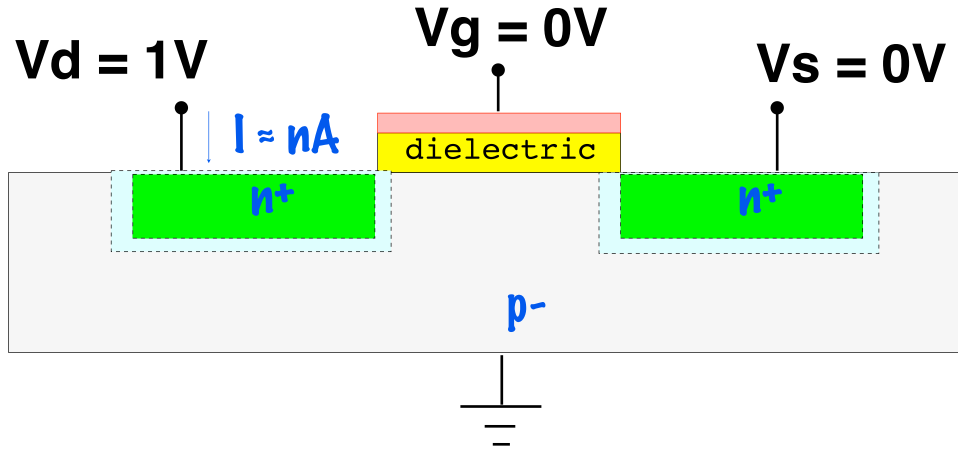
CMOS Transistors

- MOSFET (Metal Oxide Semiconductor Field Effect Transistor).



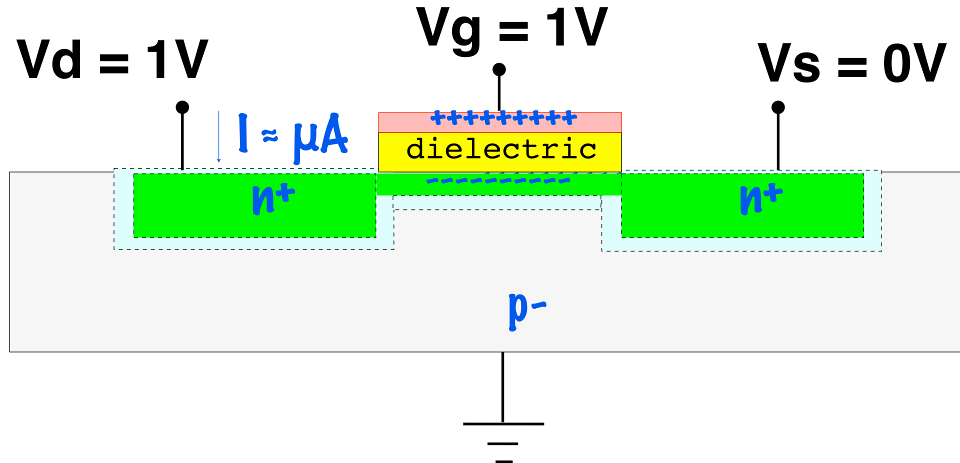
The gate acts like a capacitor. A high voltage on the gate attracts charge into the channel. If a voltage exists between the source and drain a current will flow. In its simplest approximation, the device acts like a switch.

An n-channel MOS transistor (planar)



Polysilicon gate, dielectric, and substrate form a capacitor.

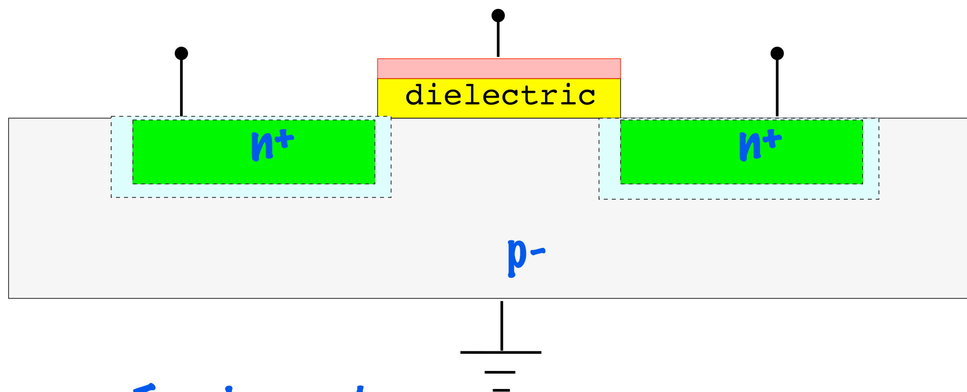
nFet is off
(I is "leakage")



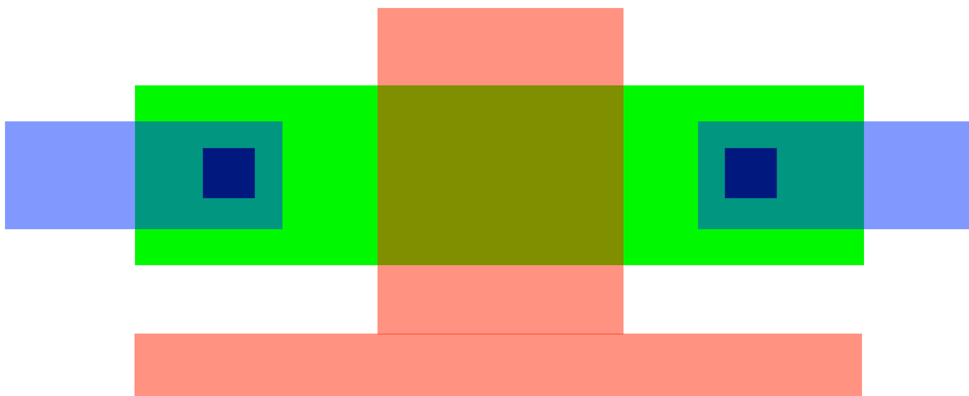
Vg = 1V, small region near the surface turns from p-type to n-type.

nFet is on.

Mask set for an n-Fet (circa 1986)



Top-down view:



Masks

#1: n+ diffusion

#2: poly (gate)

#3: diff contact

#4: metal

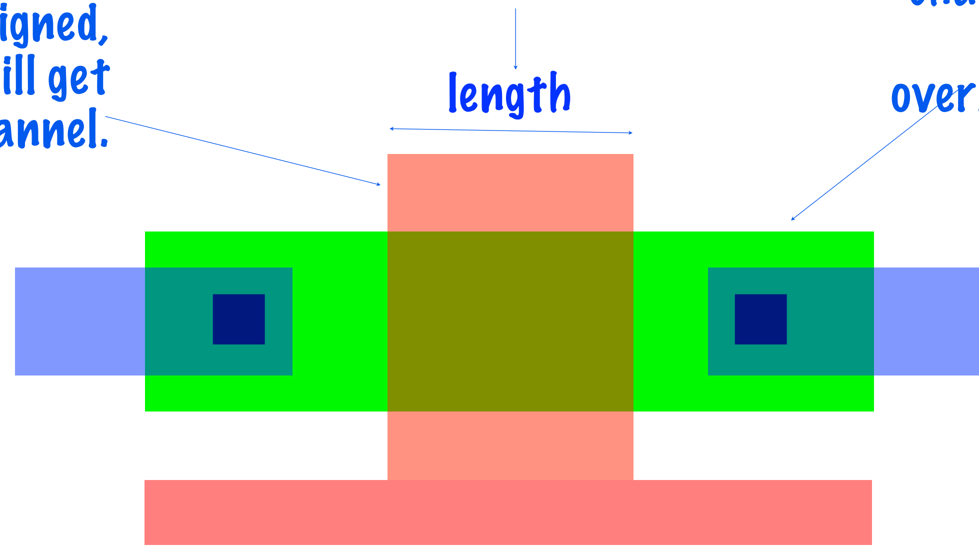
Layers to do p-Fet not shown. Modern processes have 6 to 10 metal layers (or more) (in 1986: 2).

“Design rules” for masks, 1986 ...

Poly overhang. So that if masks are misaligned, we still get channel.

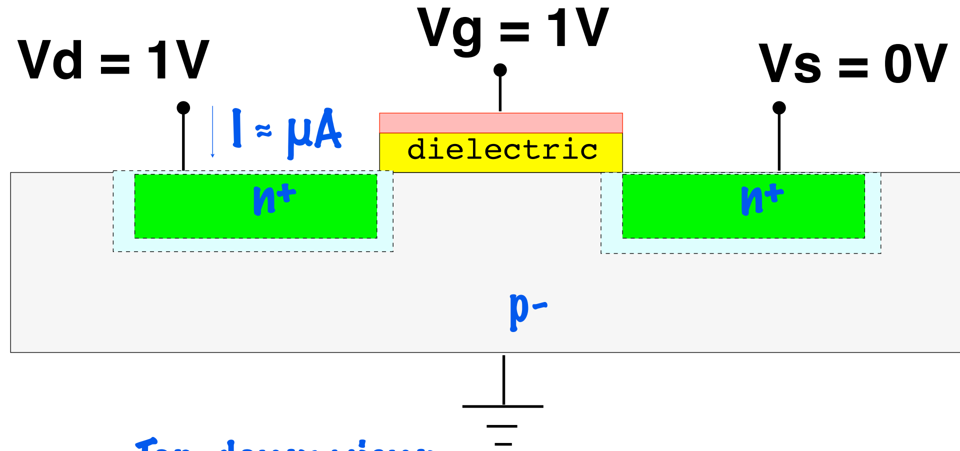
Minimum gate length. So that the source and drain depletion regions do not meet!

Metal rules: Contact separation from channel, one fixed contact size, overlap rules with metal, etc ...

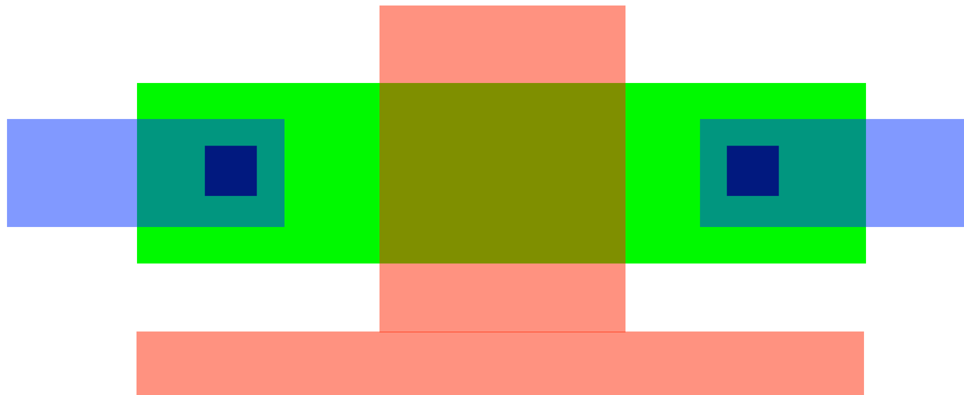


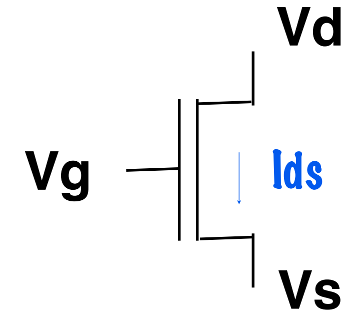
#1: n+ diffusion
#2: poly (gate)
#3: diff contact
#4: metal

How a fab uses a mask set to make an IC



Top-down view:



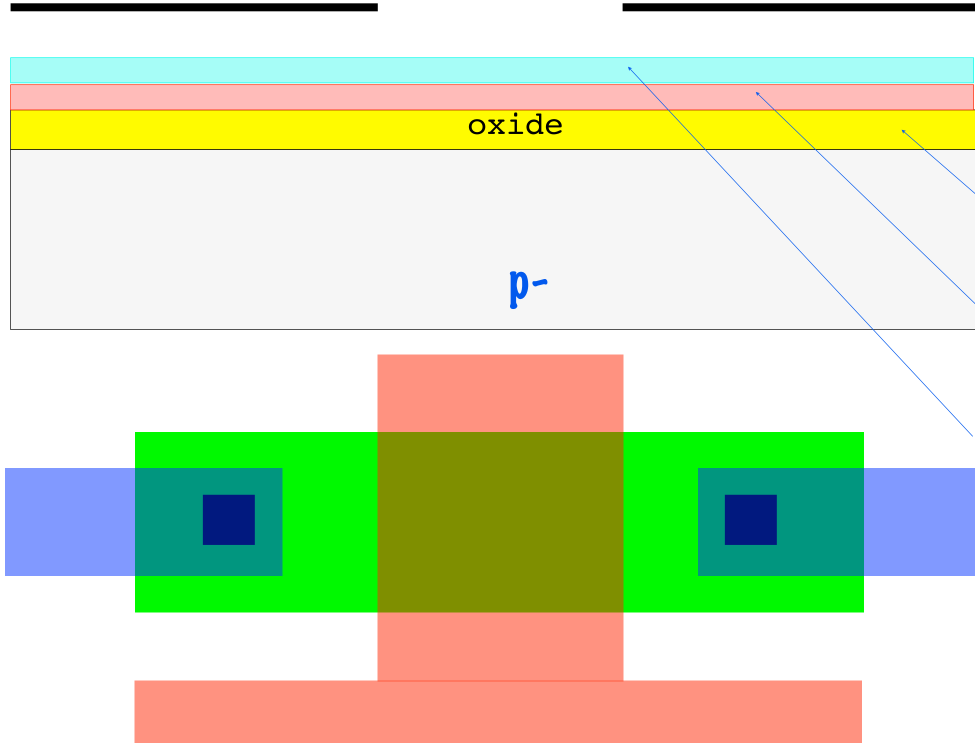


Masks

#1: n+ diffusion
#2: poly (gate)
#3: diff contact
#4: metal

Start with an un-doped wafer ...

UV hardens exposed resist. A wafer wash leaves only hard resist.



Steps

#1: dope wafer p-

#2: grow gate oxide

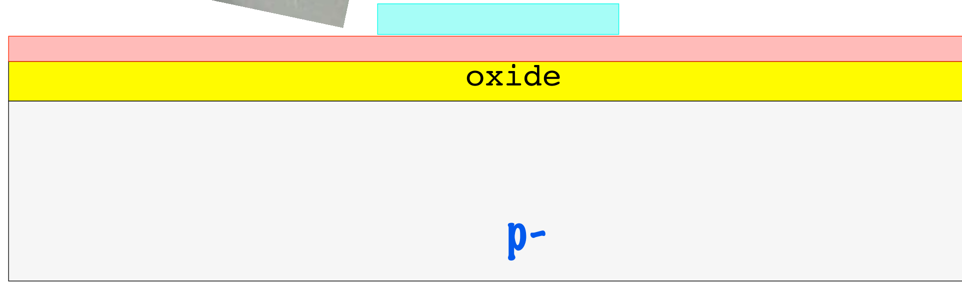
#3: deposit polysilicon

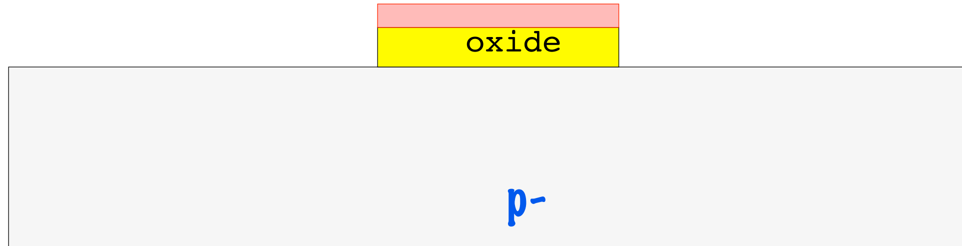
#4: spin on photoresist

#5: place positive poly mask and expose with UV.

Wet etch to remove unmasked ...

HF acid etches through poly and oxide, but not hardened resist.



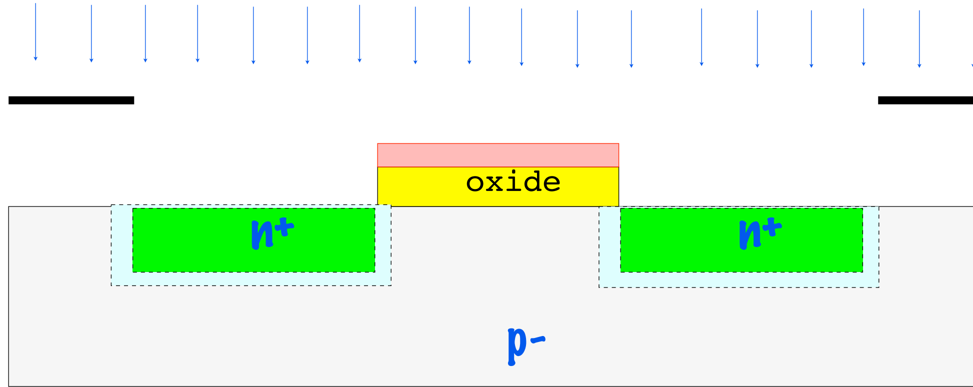


After etch and resist removal

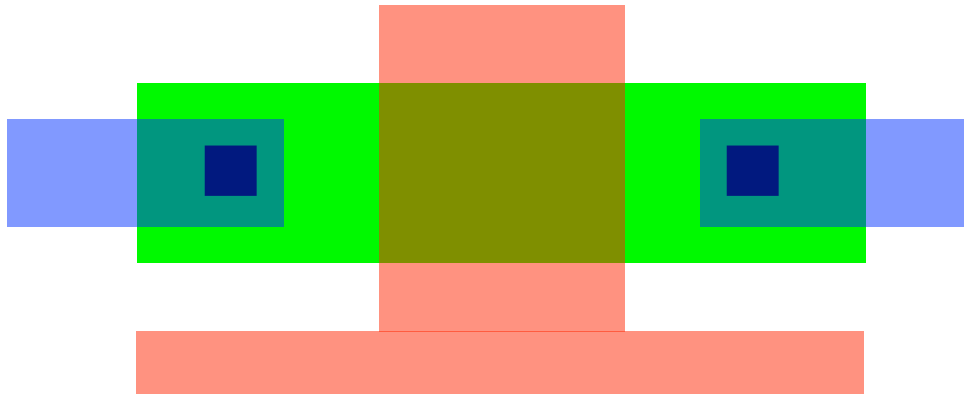
Use diffusion mask to implant n-type

accelerated donor atoms

Arsenic
As
74.922

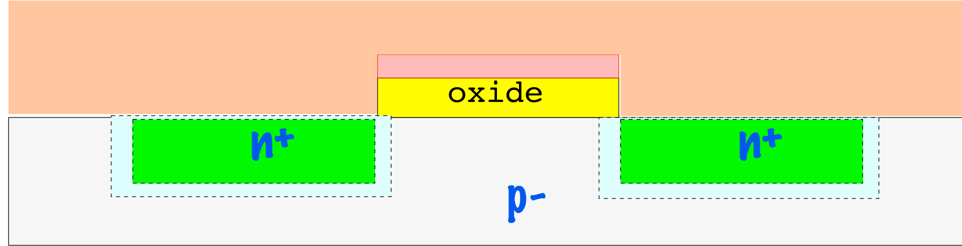


Notice how donor atoms are blocked by gate and do not enter channel.

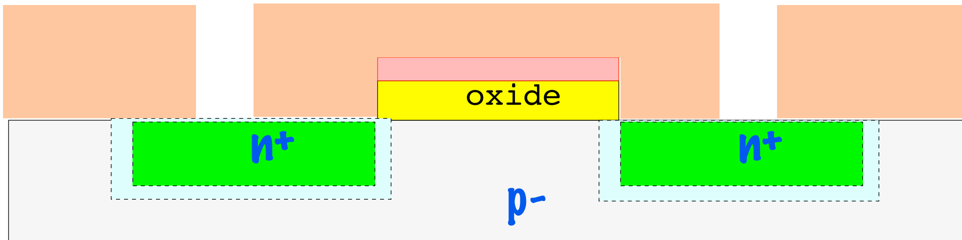


Thus, the channel is "self-aligned", precise mask alignment is not needed!

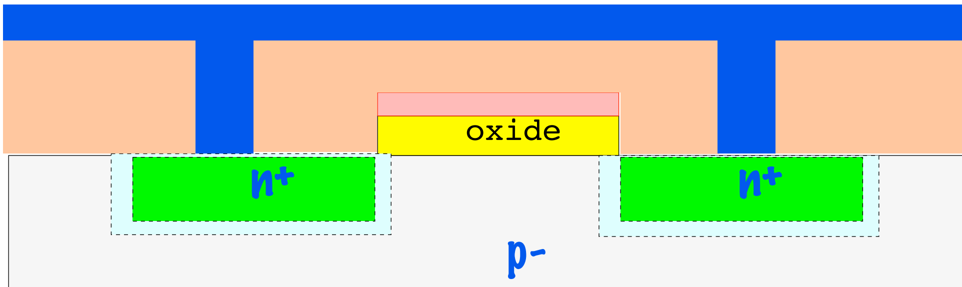
Metallization completes device



Grow a thick oxide on top of the wafer.

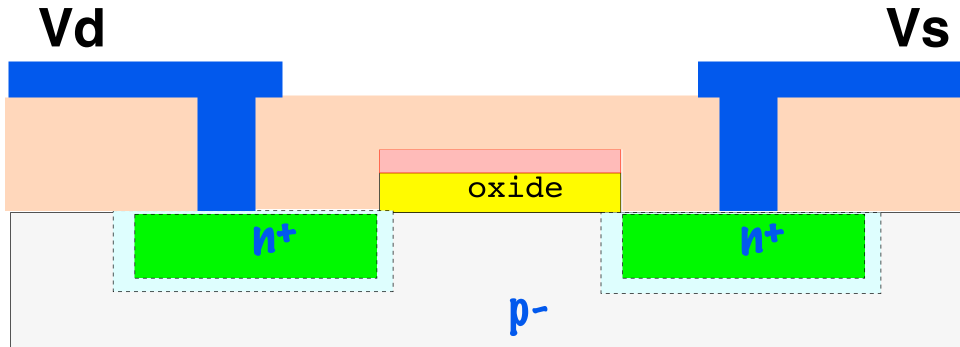


Mask and etch to make contact holes

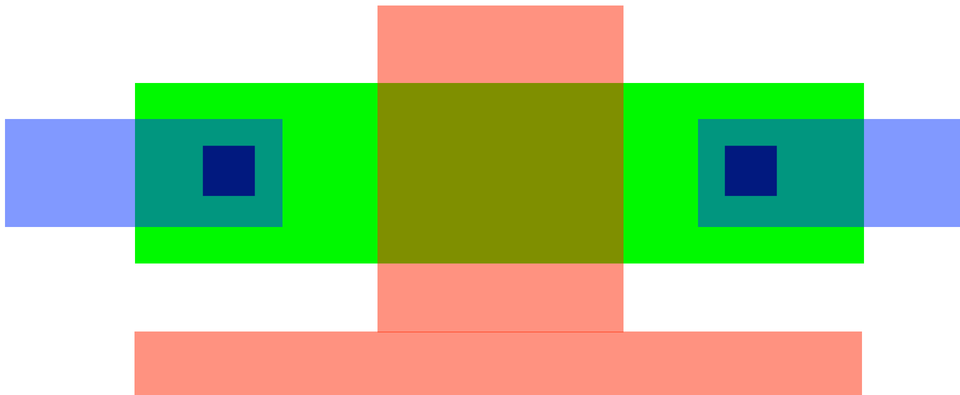


Put a layer of metal on chip. Be sure to fill in the holes!

Final product ...



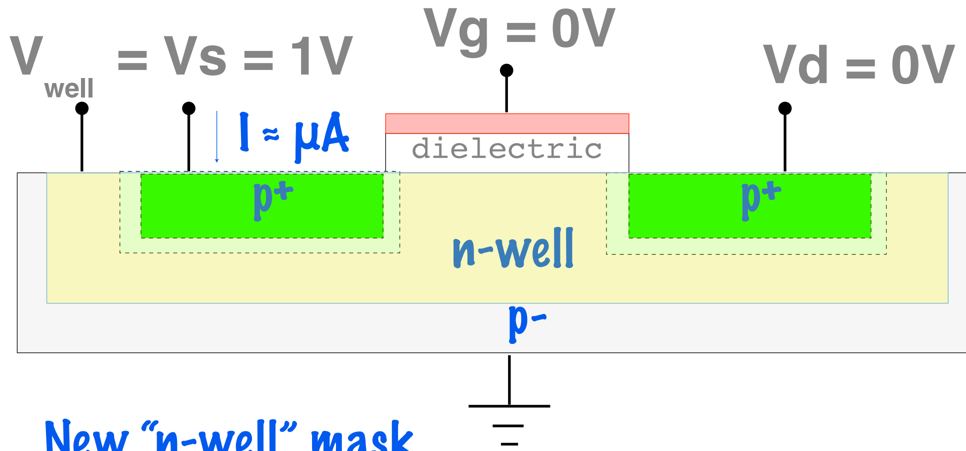
Top-down view:



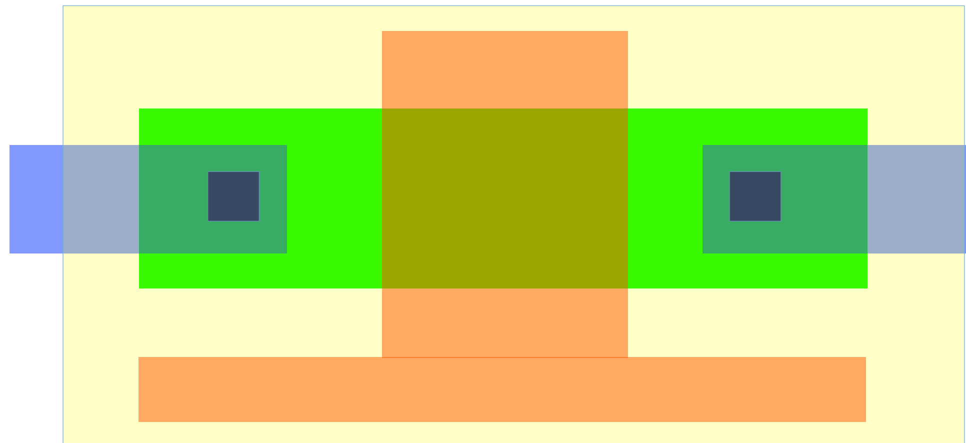
"The planar process"

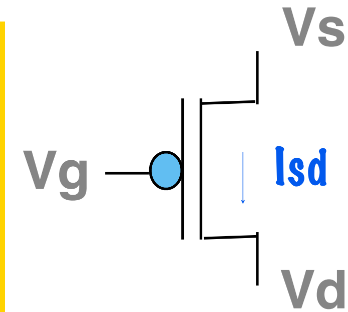
Jean Hoerni, Fairchild Semiconductor 1958

p-Fet: Change polarity of everything



New "n-well" mask

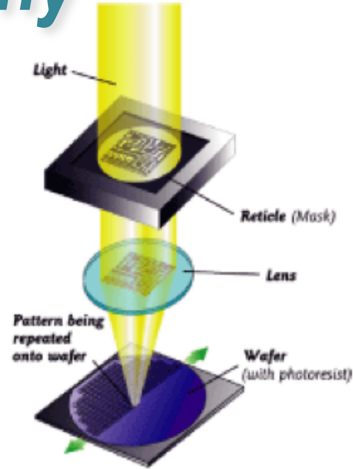




"Mobility" of holes is slower than electrons.

p-Fets drive less current than n-Fets, all else being equal

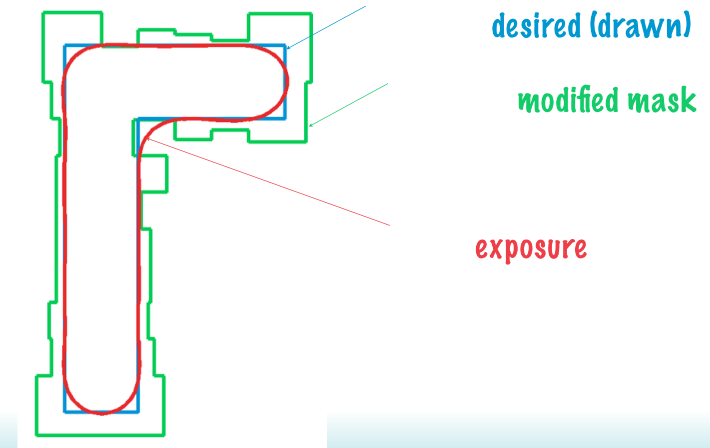
Lithography



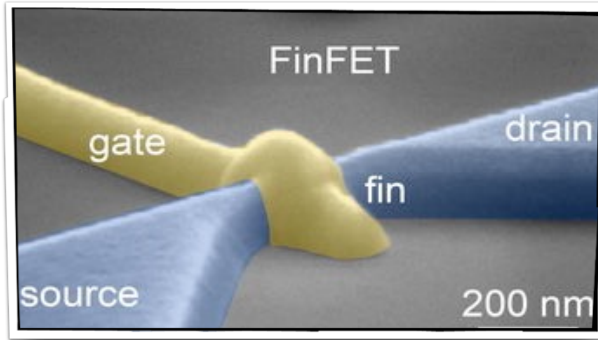
A lithography device [International Society of Optical Engineering]

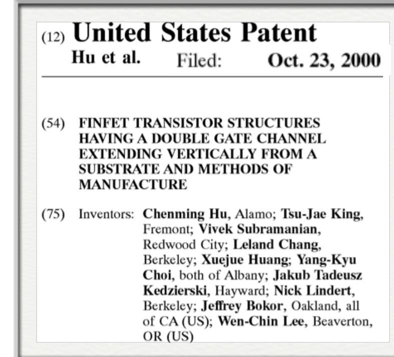
- Photolithography tools use deep ultraviolet (DUV) light with wavelengths of 248 and 193 nm, which allow minimum feature sizes below 50 nm. Newer processing uses extreme ultraviolet (EUV) with wavelength of 13.5nm.

- Optical proximity correction (OPC) is an enhancement technique commonly used to compensate for image errors due to diffraction or process effects.



Latest Modern Process



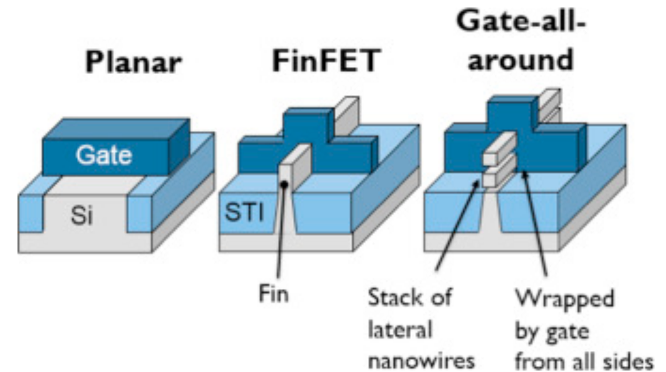

(12) **United States Patent**
Hu et al. Filed: **Oct. 23, 2000**

(54) **FINFET TRANSISTOR STRUCTURES HAVING A DOUBLE GATE CHANNEL EXTENDING VERTICALLY FROM A SUBSTRATE AND METHODS OF MANUFACTURE**

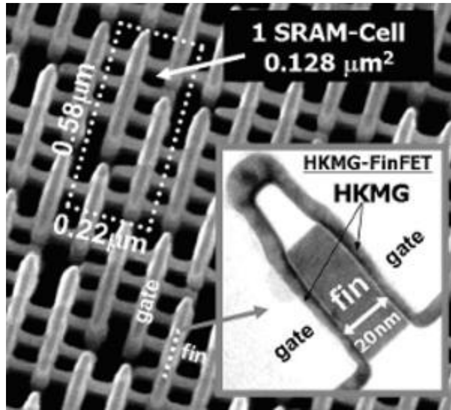
(75) Inventors: **Chenming Hu**, Alamo; **Tsu-Jae King**, Fremont; **Vivek Subramanian**, Redwood City; **Leland Chang**, Berkeley; **Xuejue Huang**; **Yang-Kyu Choi**, both of Albany; **Jakub Tadeusz Kedzierski**, Hayward; **Nick Lindert**, Berkeley; **Jeffrey Bokor**, Oakland, all of CA (US); **Wen-Chin Lee**, Beaverton, OR (US)

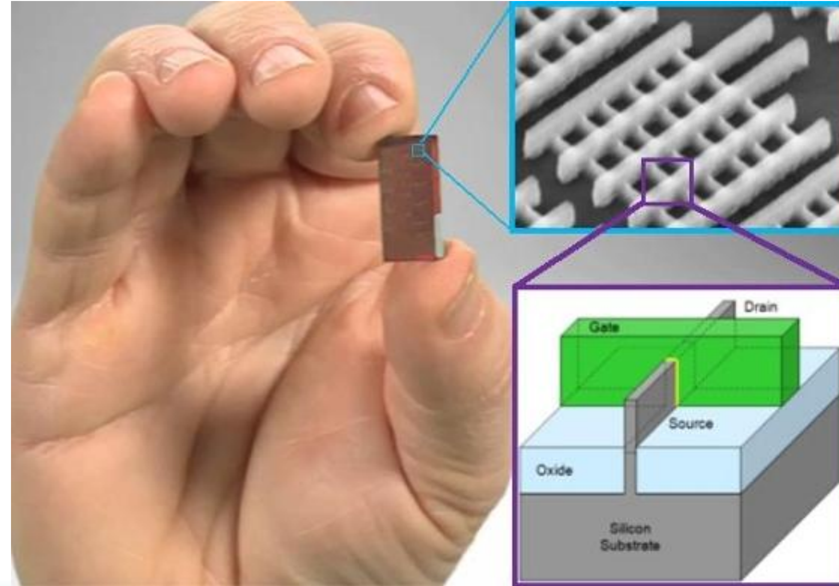
Transistor channel is a raised fin.

Gate controls channel from sides and top.



CMOS Transistors – State-of-the-Art

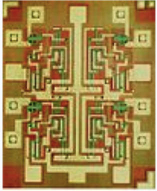




SIlicon Foundries

- Separate the designer from the fabricator: Modeled after the printing industry. (Very few authors actually own and run printing presses!)
- Standard geometric design rules are the key: these form the “contract” between the designer and manufacturer.
- Designer sends the layout (in GDS format), foundry manufactures the chip and send back. Designer promises not to violate the design rules. Foundry promises to accurately follow layout.
  - A scalable model for the industry:
    - IC fab is expensive and complex
    - Amortizes the expense over many designers (batch processing with deep queues help).
    - Designers and companies not held back by need to develop and maintain large expensive factories.
    - “fabless” semiconductor companies - lots of these and very few foundries.

Semiconductor device fabrication



MOSFET scaling (process nodes)

- 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
- 5 nm – 2020
- 3 nm – 2022

Future

- 2 nm ~ 2024

State of the art

* From Wikipedia

7nm

As of September 2018, mass production of 7 nm devices has begun. The first mainstream 7 nm mobile processor intended for mass market use, the Apple A12 Bionic, was released at their September 2018 event. Although Huawei announced its own 7 nm processor before the Apple A12 Bionic, the Kirin 980 on August 31, 2018, the Apple A12 Bionic was released for public, mass market use to consumers before the Kirin 980. Both chips are manufactured by TSMC. On July 7, 2019, AMD officially launched their Ryzen 3000 series of central processing units, based on the TSMC 7 nm process and Zen 2 microarchitecture.

5nm

In October 2019, TSMC started sampling 5nm A14 processors for Apple. In December 2019, TSMC announced an average yield of ~80%, with a peak yield per wafer of >90% for their 5nm test chips with a die size of 17.92 mm^2. In mid 2020 TSMC claimed its (N5) 5nm process offered 1.8x the density of its 7nm N7 process, with 15% speed improvement or 30% lower power consumption; an improved sub-version (N5P) was claimed to improve on N5 with +5% speed or -10% power.[19]

On October 13, 2020, Apple announced a new iPhone 12 lineup using the A14, together with the Huawei Mate 40 lineup using the HiSilicon Kirin 9000, which were the first devices to be commercialized on TSMC's 5nm node. Later, on November 10, 2020, Apple also revealed three new Mac models using the Apple M1, another 5nm chip. According to Semianalysis, the A14 processor has a transistor density of 134 million transistors per mm^2.

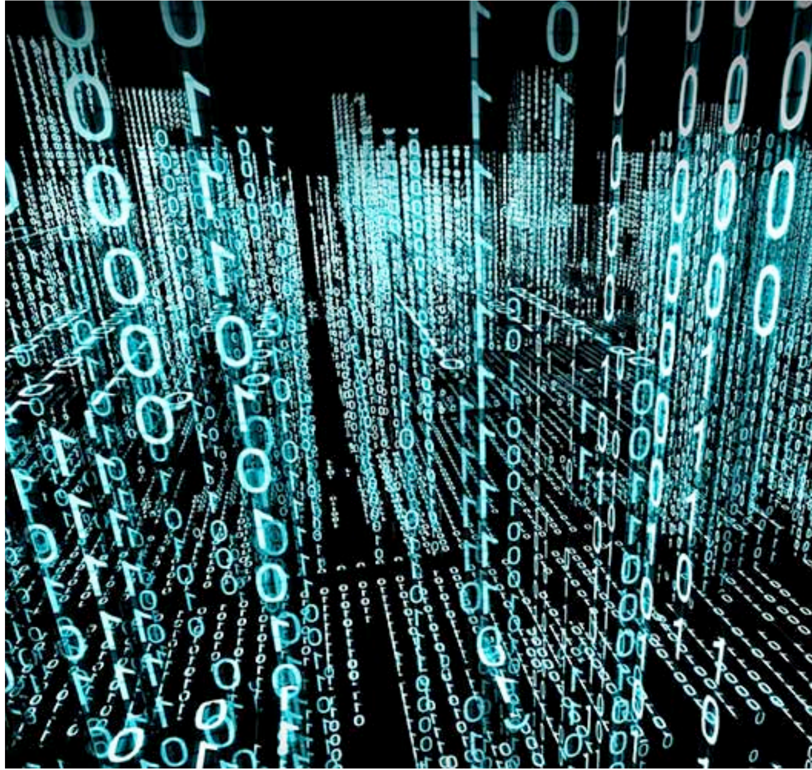
3nm

As of 2019, Intel, Samsung, and TSMC have all announced plans to put a 3 nm semiconductor node into commercial production. Samsung's 3 nm process is based on GAAFET (gate-all-around field-effect transistor) technology, a type of multi-gate MOSFET technology, while TSMC's 3nm process will still use FinFET (fin field-effect transistor) technology,[1] despite TSMC developing GAAFET transistors.

What does “3nm” mean?

- The term "3 nanometer" has no relation to any actual physical feature (such as gate length, metal pitch or gate pitch) of the transistors. According to the projections contained in the 2021 update of the International Roadmap for Devices and Systems published by IEEE Standards Association Industry Connection, a 3 nm node is expected to have a contacted gate pitch of 48 nanometers and a tightest metal pitch of 24 nanometers.
- However, in real world commercial practice, "3 nm" is used primarily as a marketing term by individual microchip manufacturers to refer to a new, improved generation of silicon semiconductor chips in terms of increased transistor density (i.e. a higher degree of miniaturization), increased speed and reduced power consumption. Moreover, there is no industry-wide agreement among different manufacturers about what numbers would define a 3 nm node.

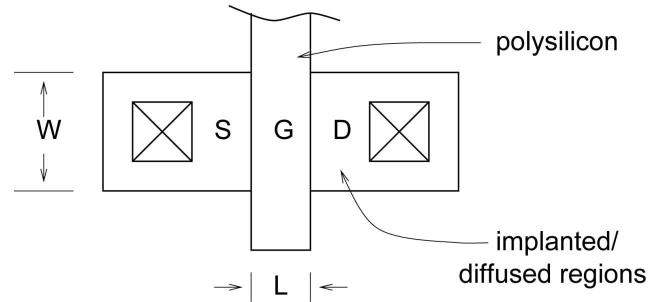
From Wikipedia



CMOS Transistors

Transistor Strength and Symmetry

1. Transistor "strength" proportional to W/L. In digital circuits, L is almost always minimal allowed by process.

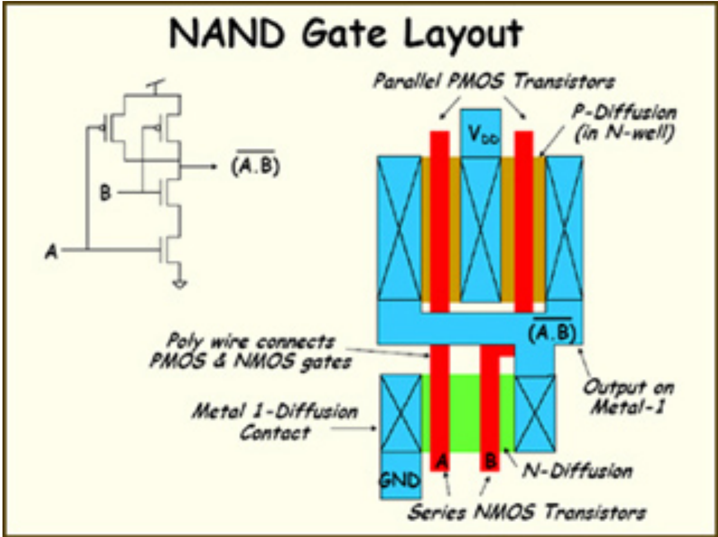


2. MOS transistors are symmetrical devices (Source and drain are interchangeable). But usually designed to be used in one direction.

For nFET, source is the node w/ the lowest voltage. For pFET source is node with highest voltage.

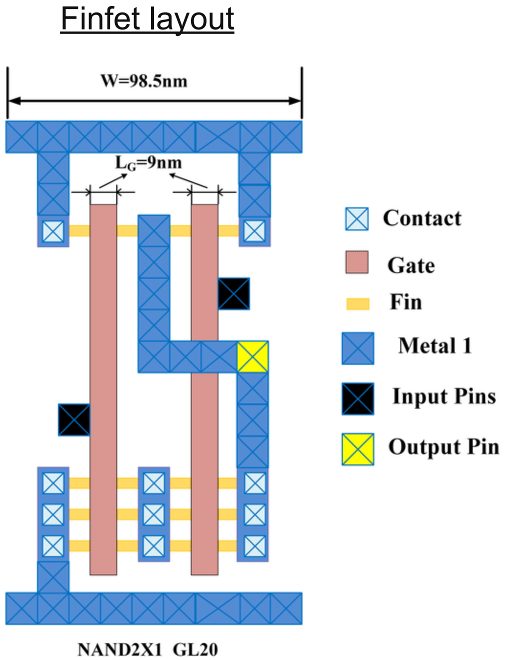
Circuit Layout Examples

- 2-input NAND



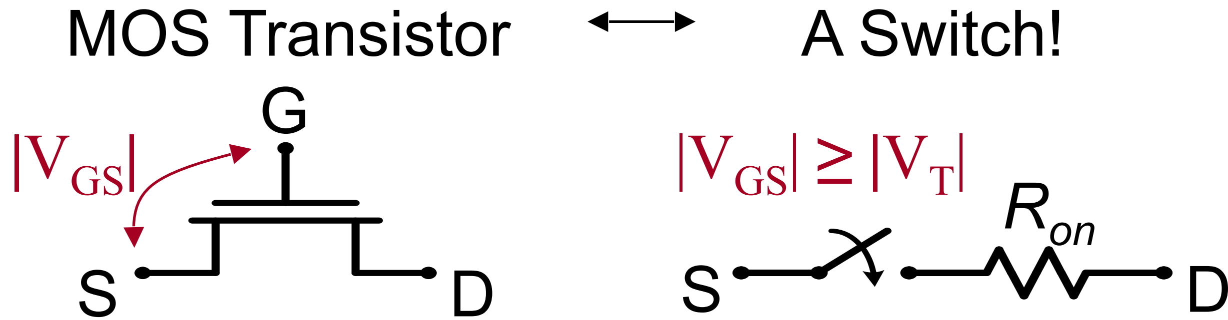
NAND gate layout from Lecture 3: CMOS Technology and Logic Gates. (Image by Professors Arvind and Asanovic.)

Finfet layout



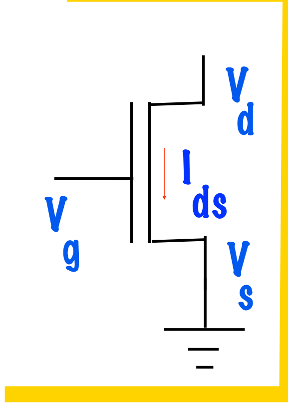
from Ji Li

MOS Transistor as a Resistive Switch



Let's look beneath the abstraction:
origins of $\boldsymbol{V_T}$ and $\boldsymbol{R_{on}}$

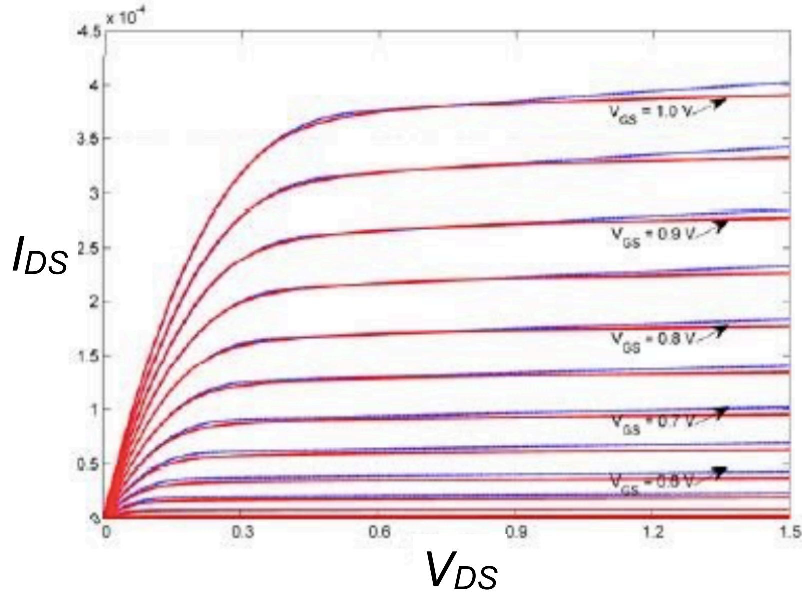
MOSFET Threshold Voltage

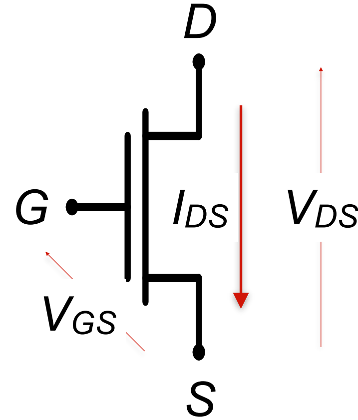


Transistor "turns on" when V_{gs} is > V_t.

Transistor "resistance"

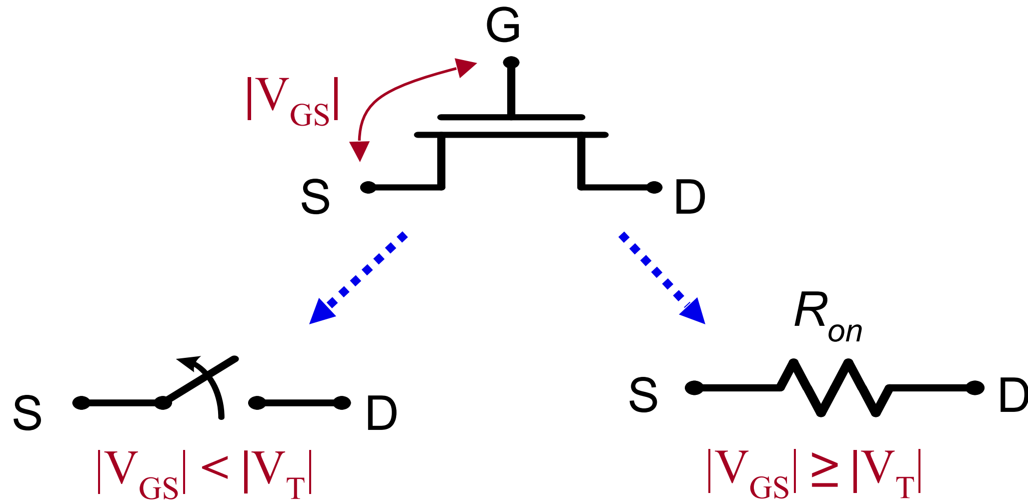
- Nonlinear I/V characteristic:



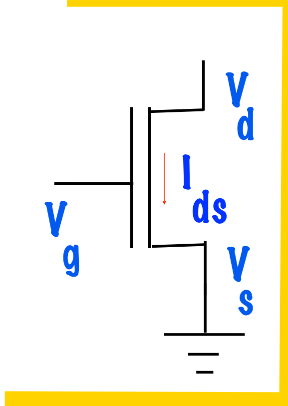


- But, linearizing makes all delay and power calculations simple (usually just 1st order ODEs):

ON/OFF Switch Model of MOS Transistor



Plot on a "Log" Scale to See "Off" Current

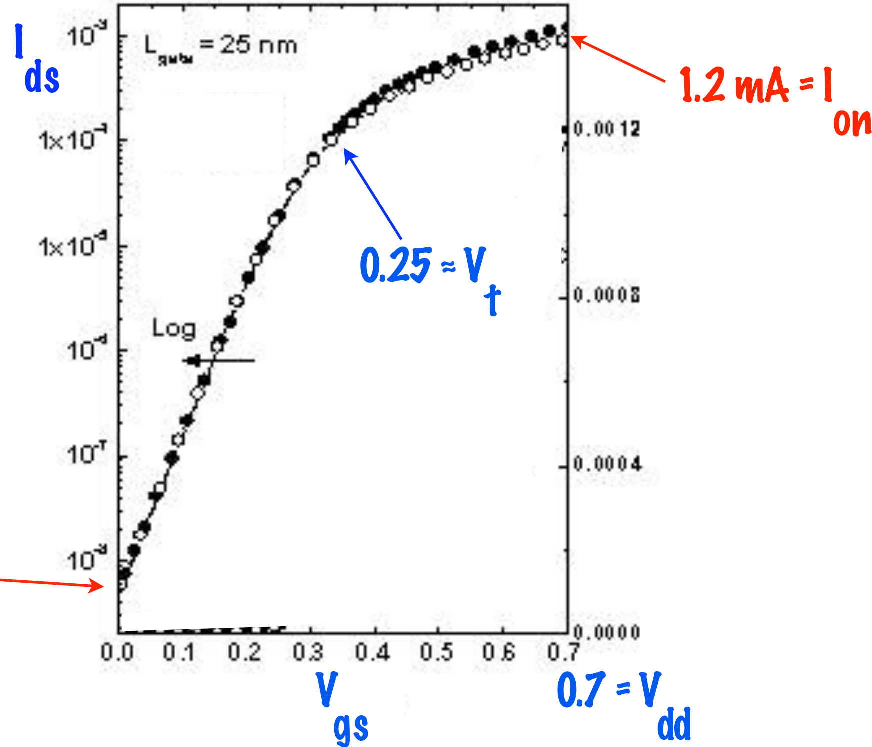


$V_{ds} = V_{dd}$

Process engineers can:

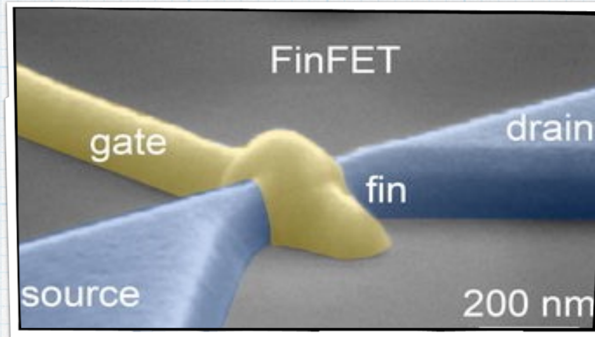
increase I_{on} by lowering V_t - but that raises I_{off}

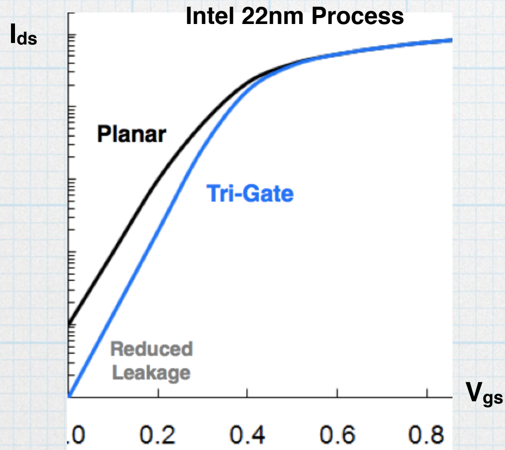
decrease I_{off} by raising V_t - but that lowers I_{on}.



I_{off} ≈ 10 nA

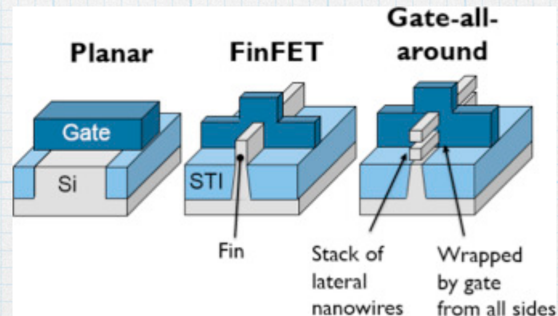
Latest Modern Process



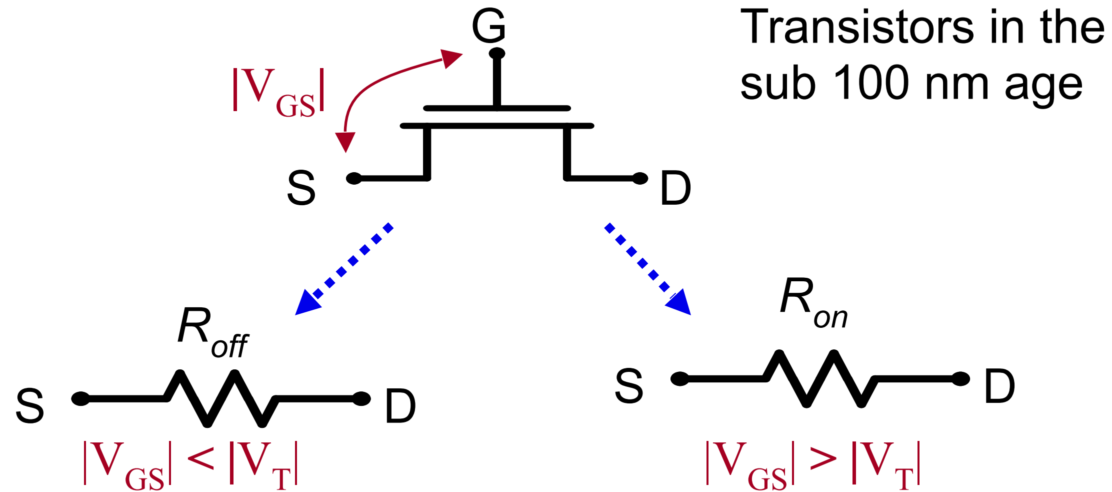


Transistor channel is a raised fin.

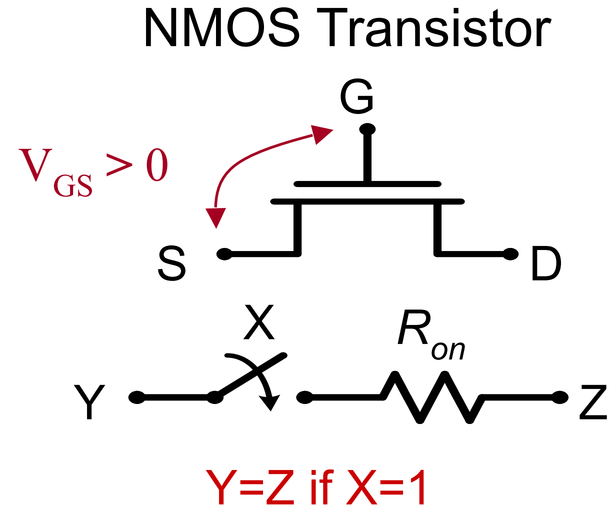
Gate controls channel from sides and top.



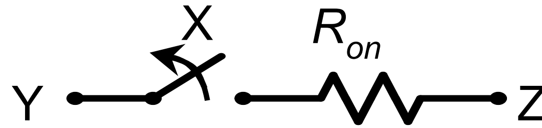
A More Realistic Switch



A Logic Perspective

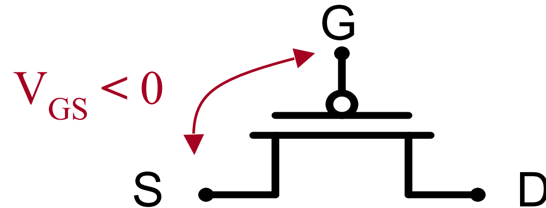


A Complementary Switch



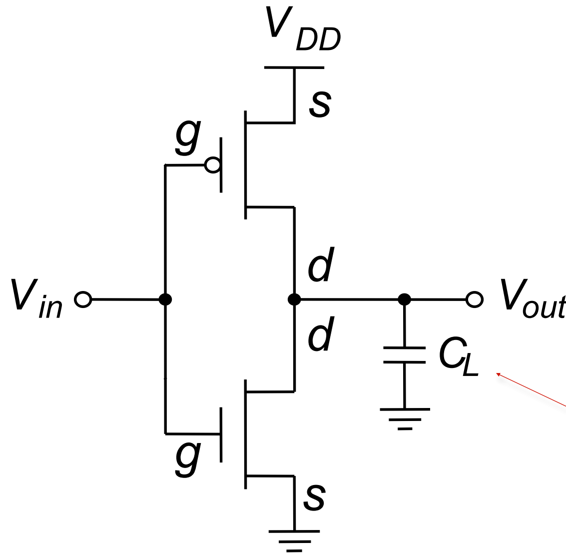
Y=Z if X=0

PMOS Transistor



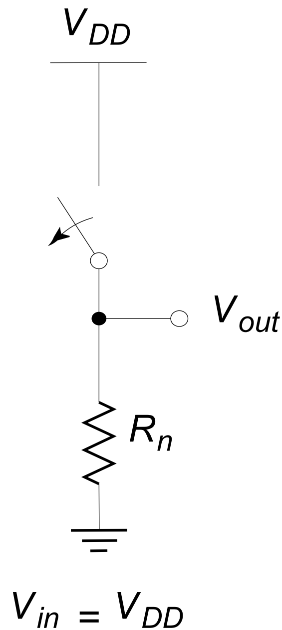
Remember, source is the node w/ the highest voltage.

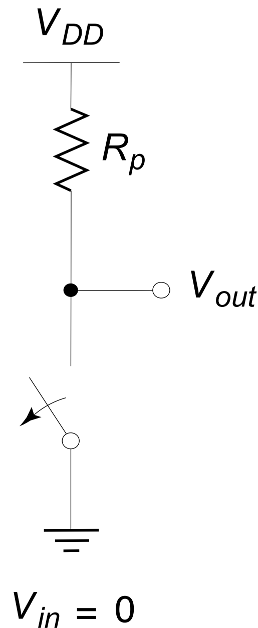
The CMOS Inverter: A First Glance



Represents the sum of all the capacitance at the output of the inverter and everything to which it connects: (drains, interconnections gate capacitance of next gate(s))

The Switch Inverter
First-Order DC Analysis*





$V_{OL} = 0$
$V_{OH} = V_{DD}$

First-order means we will ignore Capacitance.