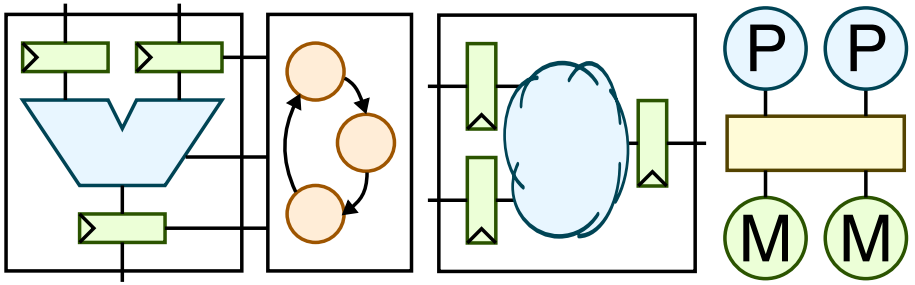
ECE 5745 Complex Digital ASIC Design
Topic 2: CMOS Devices

Christopher Batten

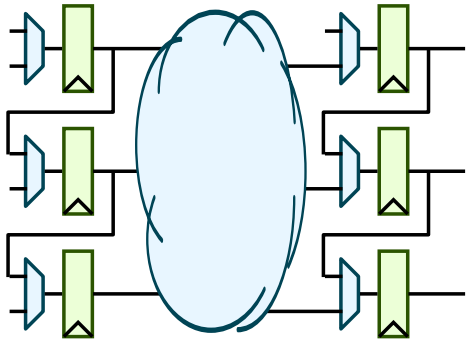
School of Electrical and Computer Engineering
Cornell University

http://www.csl.cornell.edu/courses/ece5745

Part 1: ASIC Design Overview



Topic 1
Hardware Description Languages

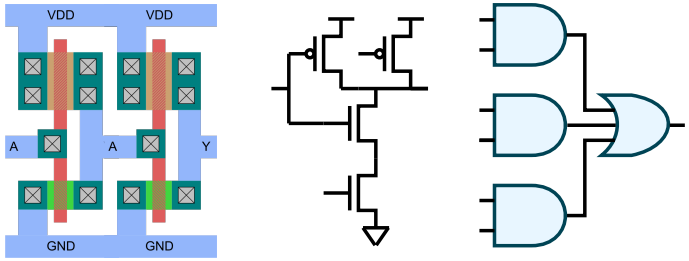


Topic 8
Testing and Verification

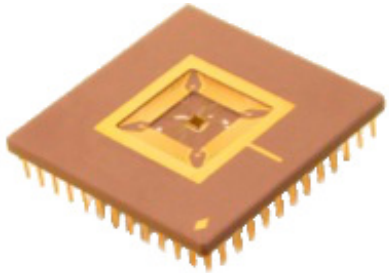
Topic 4
Full-Custom Design Methodology

Topic 6
Closing the Gap

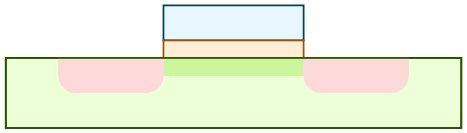
Topic 5
Automated Design Methodologies



Topic 3
CMOS Circuits



Topic 7
Clocking, Power Distribution, Packaging, and I/O



Topic 2
CMOS Devices

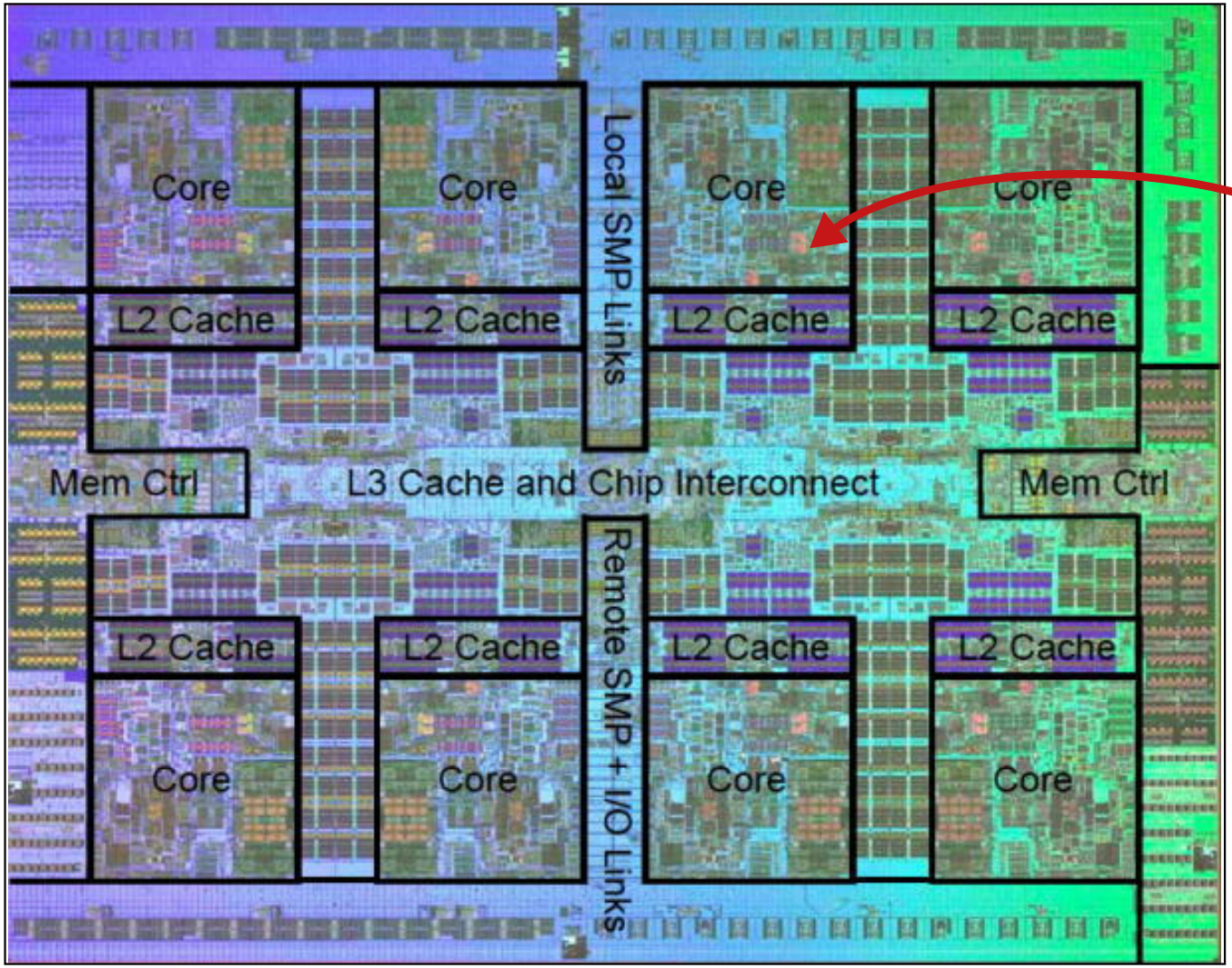
Agenda

Simple Transistor RC Model

Simple Wire RC Model

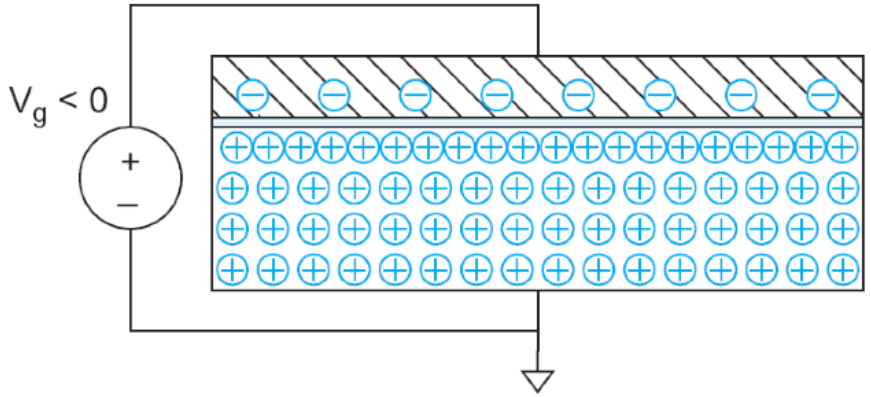
CMOS Fabrication

Fundamental Building Block: MOSFET Transistor



IBM Power 7
1.2 Billion
Transistors

"Metal"-Oxide-Semiconductor Structure

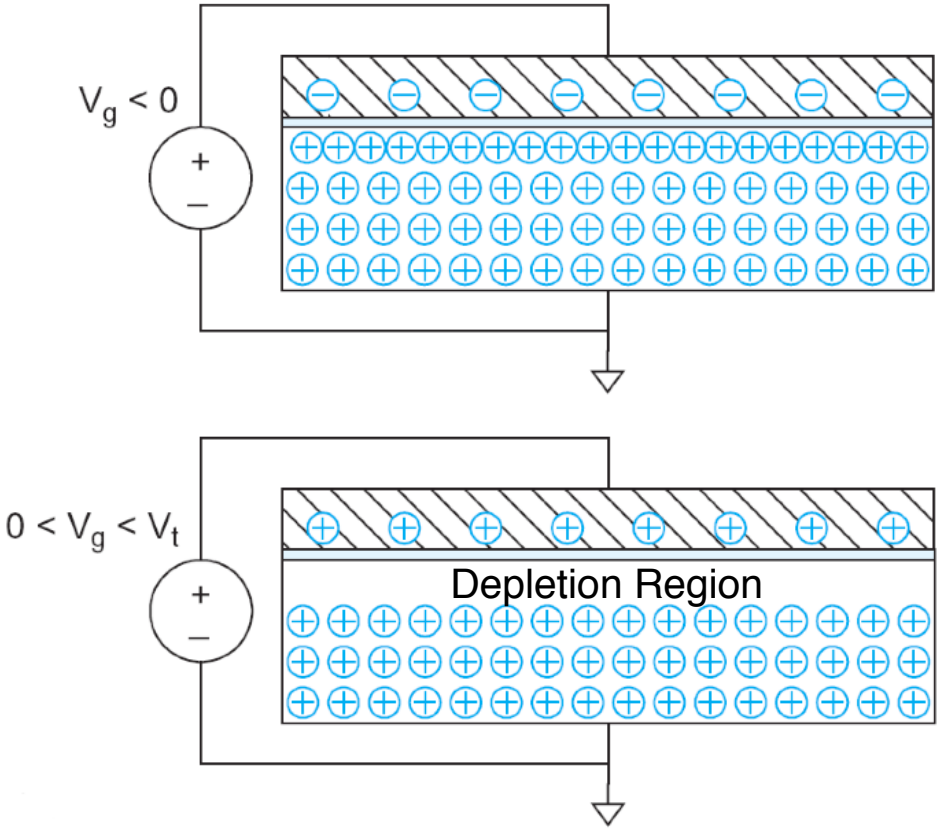


Polysilicon Gate
Silicon Dioxide Insulator
p-Type Silicon Body
(doped to create mobile majority carriers, positively charged holes)

Accumulation: Battery puts negative charge on gate, attracts positively-charged majority carriers in p-type silicon body

Adapted from [Weste'11]

"Metal"-Oxide-Semiconductor Structure

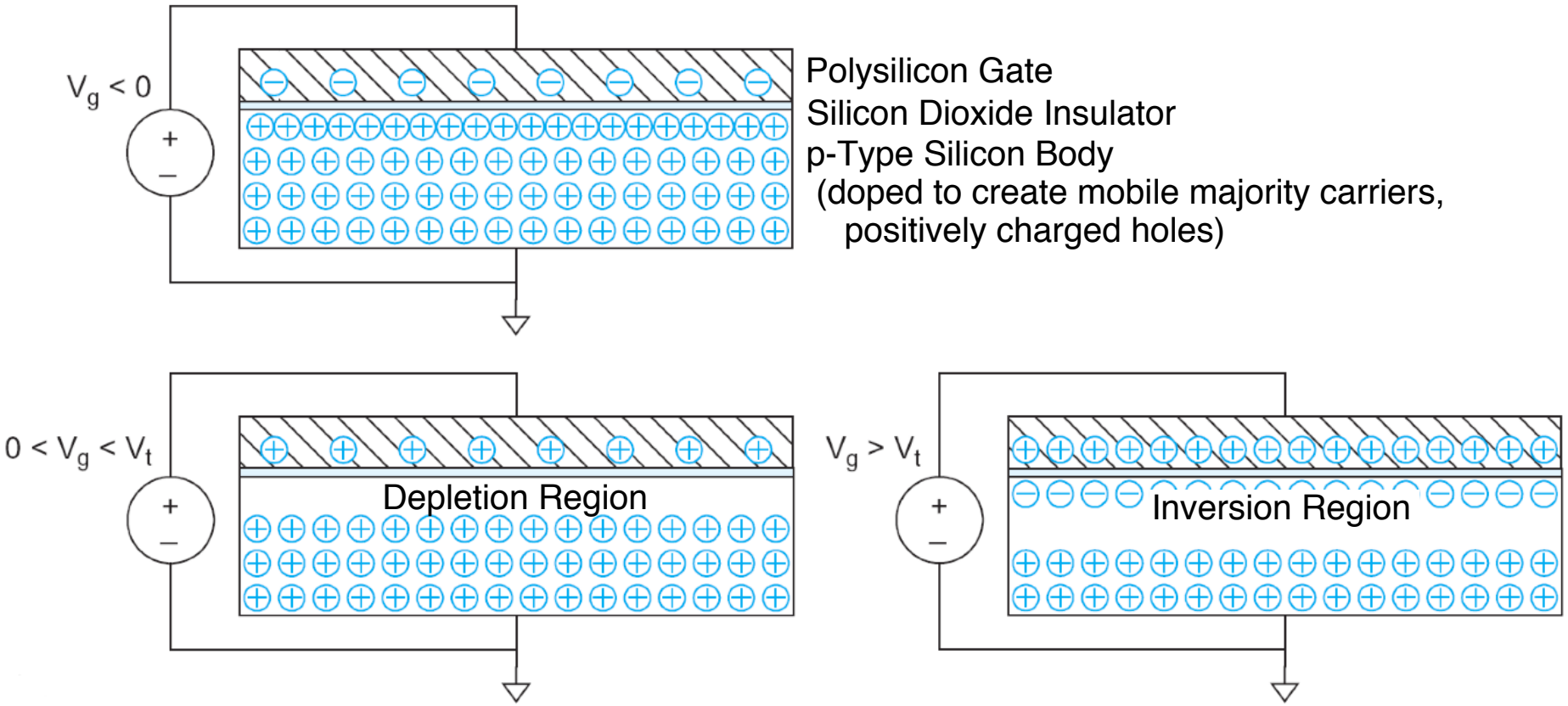


Polysilicon Gate
Silicon Dioxide Insulator
p-Type Silicon Body
(doped to create mobile majority carriers, positively charged holes)

Depletion: Battery puts positive charge on gate, pushes positively-charged carriers away from surface, uncovers some negatively-charged dopant atoms in substrate

Adapted from [Weste'11]

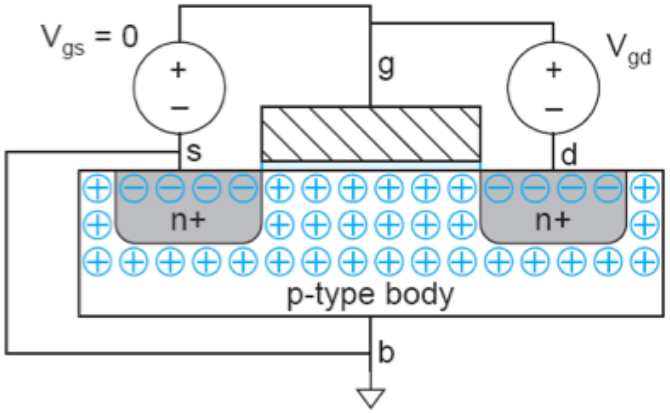
"Metal"-Oxide-Semiconductor Structure



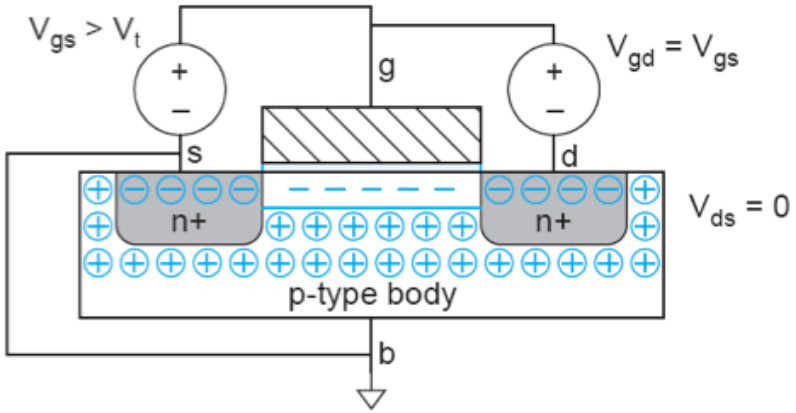
Inversion: Battery puts more positive charge on gate, instead of pushing holes even further away, draws free electrons to surface. Where did electrons come from? No electron donors in p-type silicon; electron/hole pairs always being generated by thermal excitation – electrons caught by efield in depletion region

Adapted from [Weste'11]

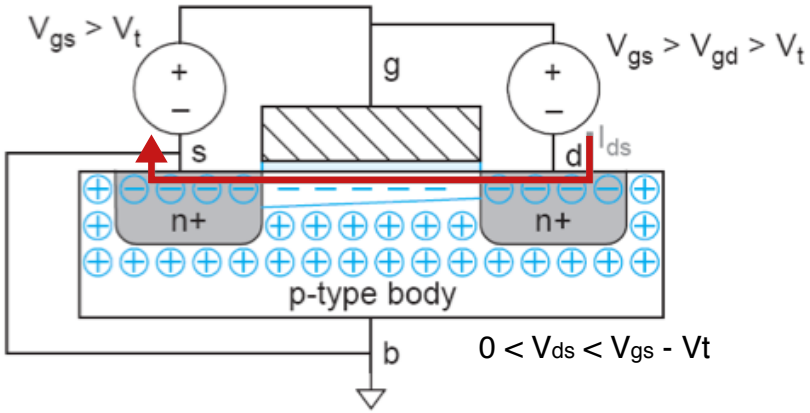
NMOS Transistor



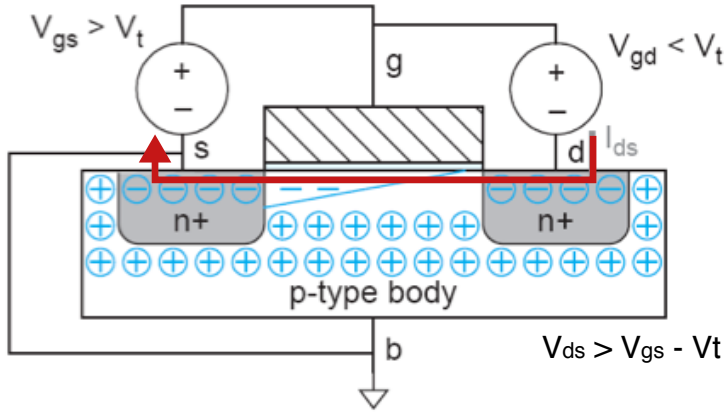
Cutoff: V_{gs} = 0V, Vds can be 0V or Vdd
No Channel, I_{ds} = 0



Linear: V_{gs} = Vdd, V_{ds} = 0V
Channel Formed, I_{ds} increases with V_{ds}



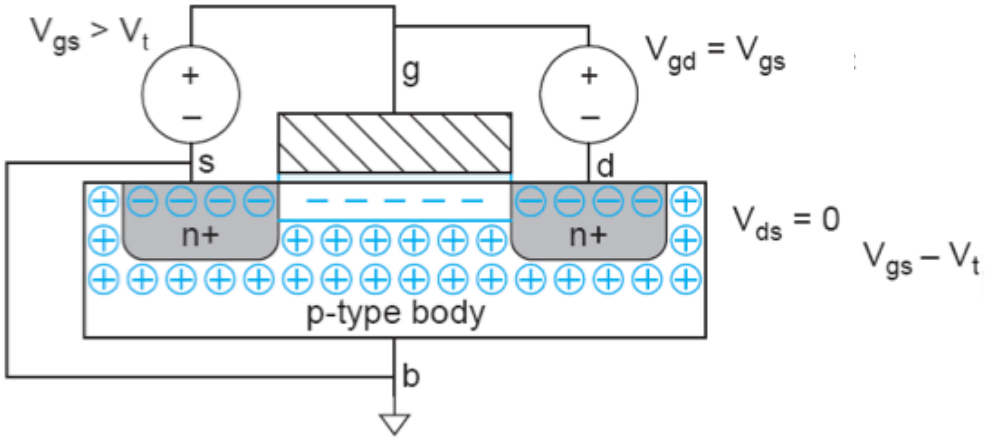
Linear: Vgs = Vdd, Vds = Vdd
Channel Formed, I_{ds} increases with V_{ds}

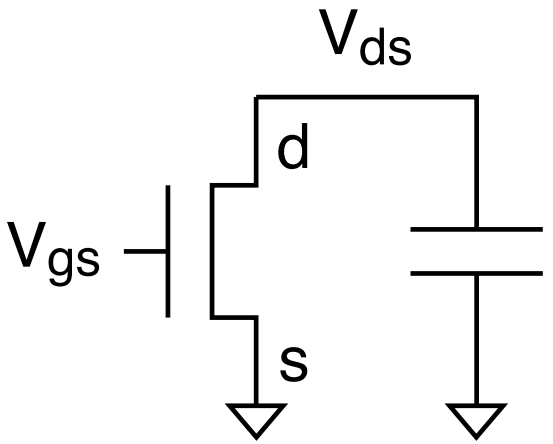


Saturation: Channel Pinched Off,
I_{ds} independent of V_{ds}

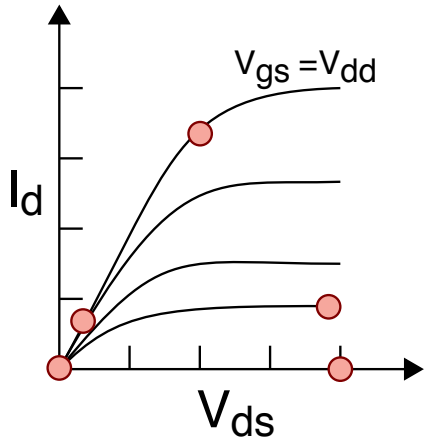
Adapted from [Weste'11]

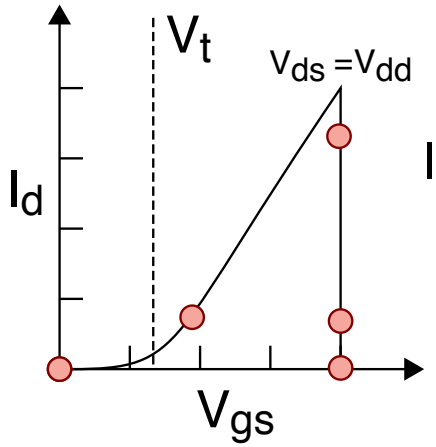
Simple NMOS Circuit

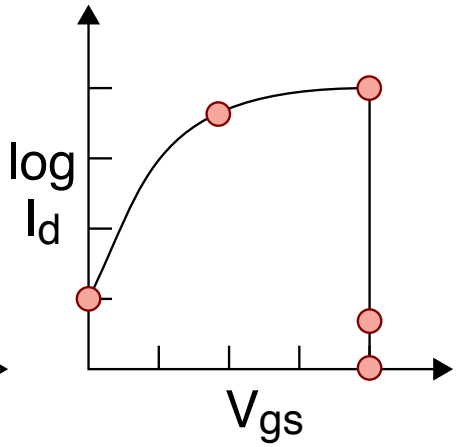


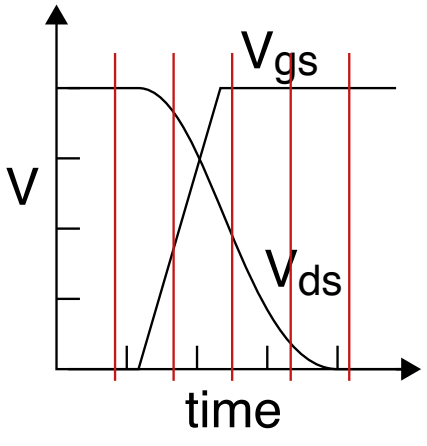


Linear: Channel Formed
I_{ds} increases with V_{ds}

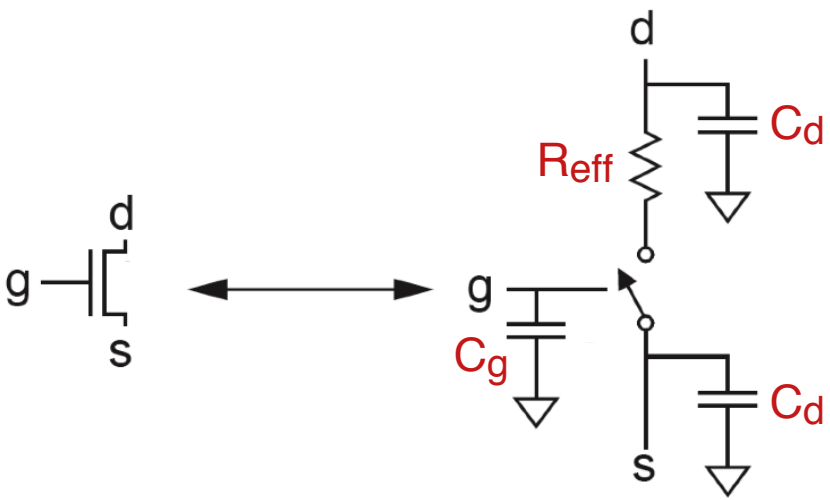


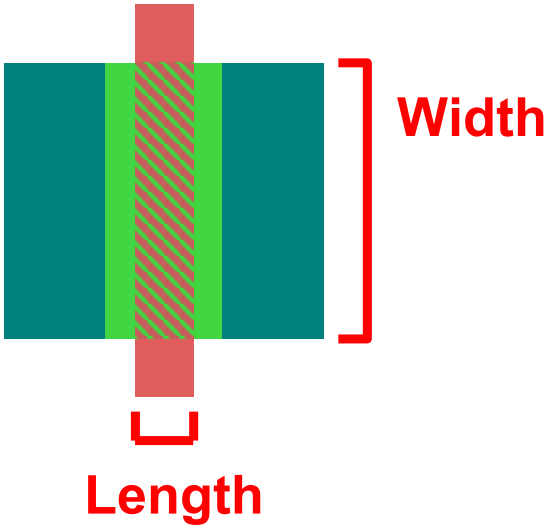






Key Qualitative Characteristics of MOSFET Transistors





- V_t sets when transistor turns on, impacts leakage current
- $I_d \propto \mu \times (W/L)$
- $\mu_n > \mu_p \implies R_{N,eff} < R_{P,eff}$
- $C_g \propto (W \times L)$
- $C_d \propto W$
- $\uparrow W = \downarrow R_{eff} = \uparrow I_d = \uparrow C_d, C_g$
- $\uparrow L = \uparrow R_{eff} = \downarrow I_d = \uparrow C_g$

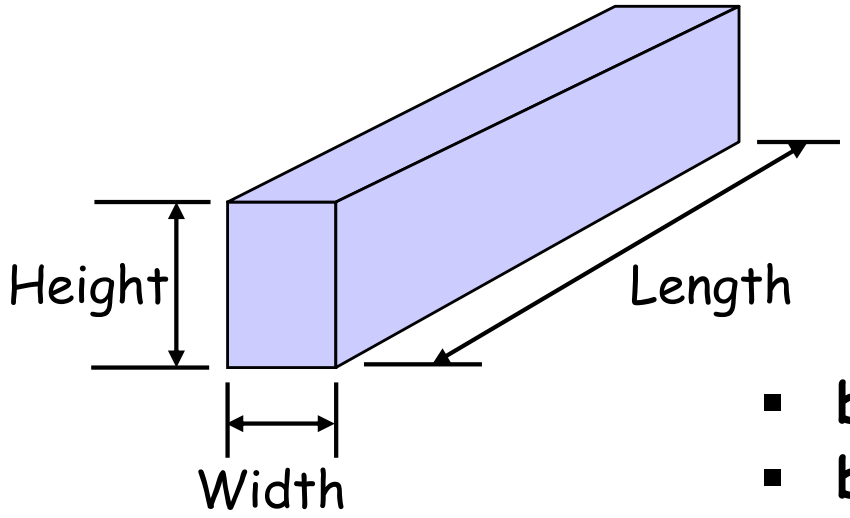
Agenda

Simple Transistor RC Model

Simple Wire RC Model

CMOS Fabrication

Wire Resistance



$$\text{resistance} = \frac{(\text{length} \times \text{resistivity})}{(\text{height} \times \text{width})}$$

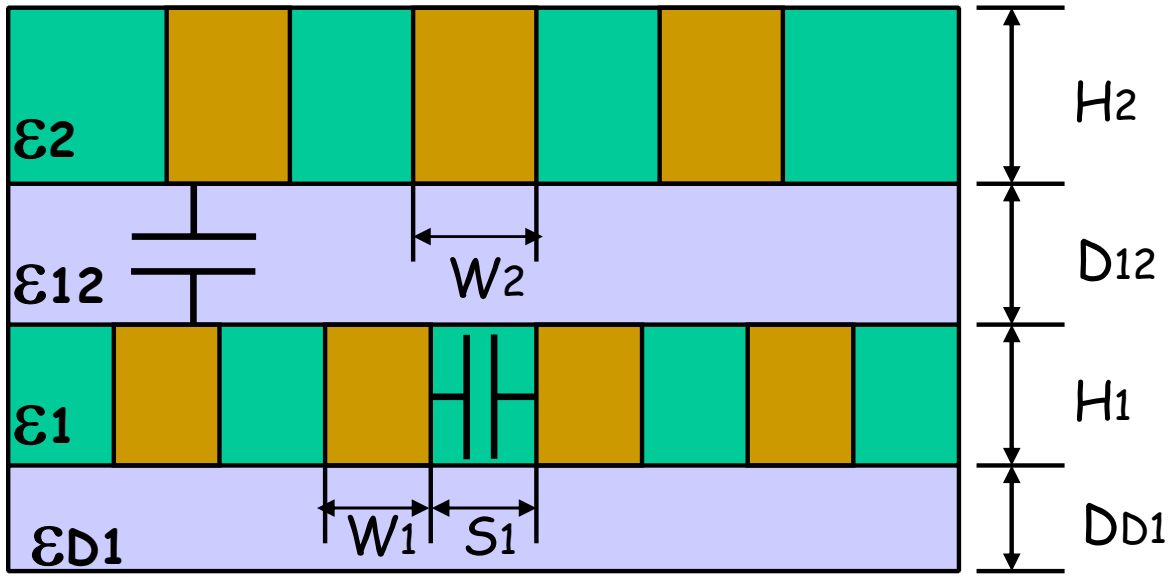
- bulk aluminum 2.8×10^{-8} Ω-m
- bulk copper 1.7×10^{-8} Ω-m
- bulk silver 1.6×10^{-8} Ω-m

- Thickness fixed in given manufacturing process
- Resistances quoted as Ω/square
- TSMC 0.18μm 6 Aluminum metal layers
  - M1-5 0.08 Ω/square (0.5 μm x 1mm wire = 160 Ω)
  - M6 0.03 Ω/square (0.5 μm x 1mm wire = 60 Ω)

R_{sq} = resistivity / height resistance = R_{sq} × (length / width)

Adapted from [Terman'02]

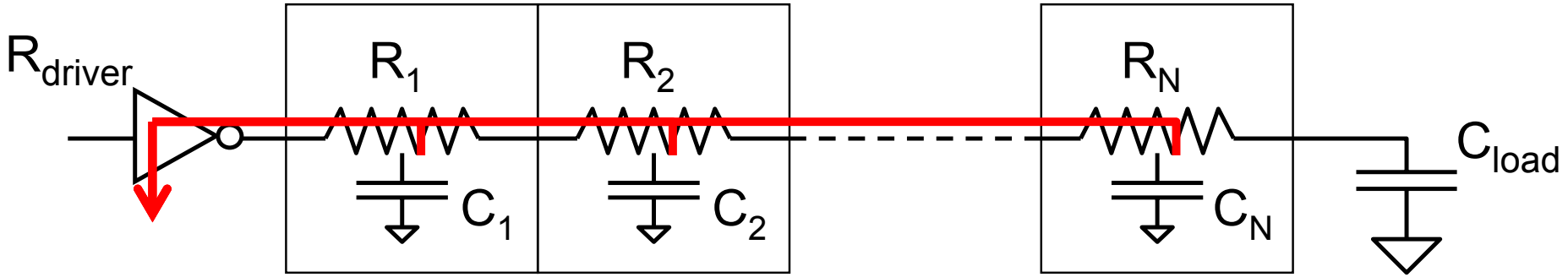
Wire Capacitance



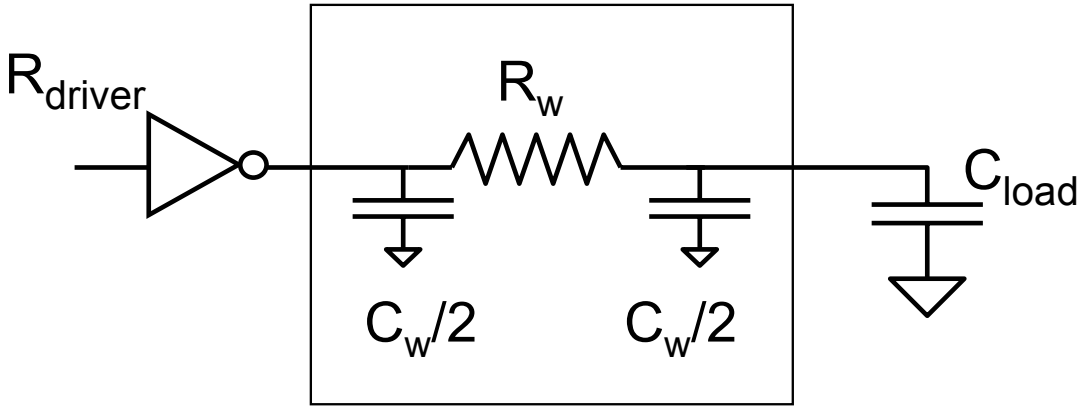
- Capacitance depends on geometry of surrounding wires and relative permittivity, ε_r,of dielectric
  - silicon dioxide, SiO_2 $\varepsilon_r = 3.9$
  - silicon flouride, SiOF $\varepsilon_r = 3.1$
  - SiLK™ polymer, $\varepsilon_r = 2.6$

Adapted from [Terman'02]

Key Qualitative Characteristics of Wires



Because both wire resistance and wire capacitance increase with length, wire delay grows quadratically with length



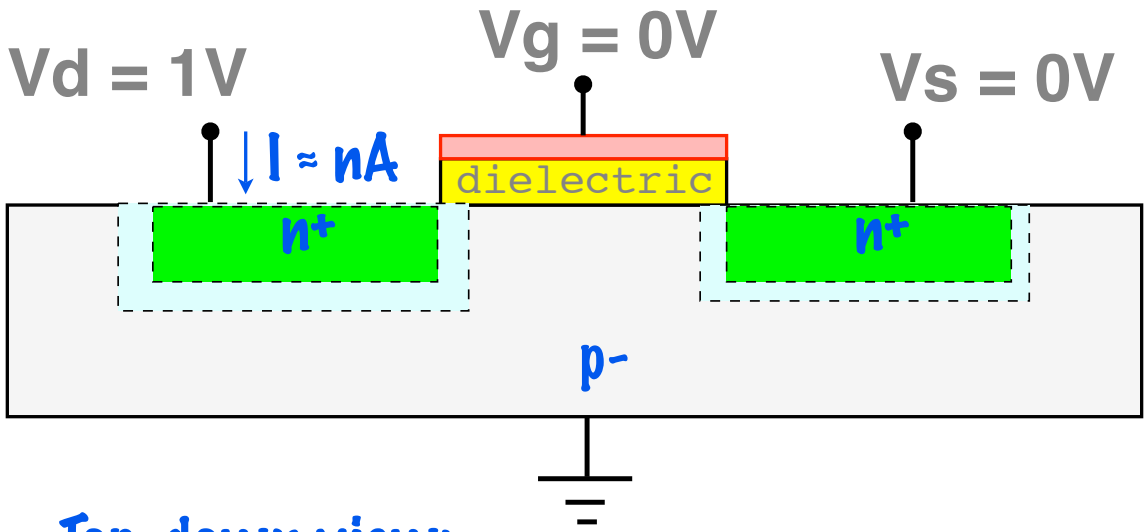
Agenda

Simple Transistor RC Model

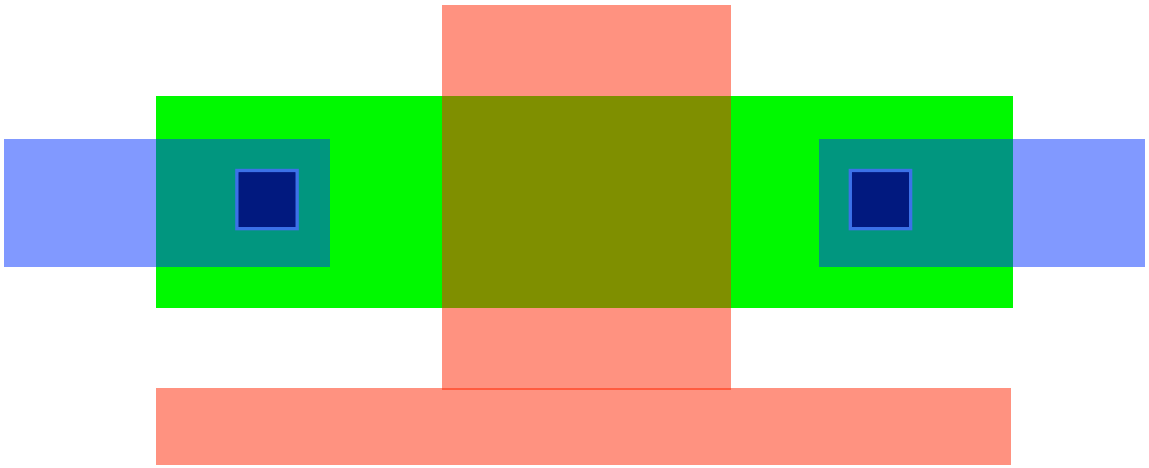
Simple Wire RC Model

CMOS Fabrication

Mask Set for NMOS Transistor (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).

38

Adapted from [Asanovic'11]

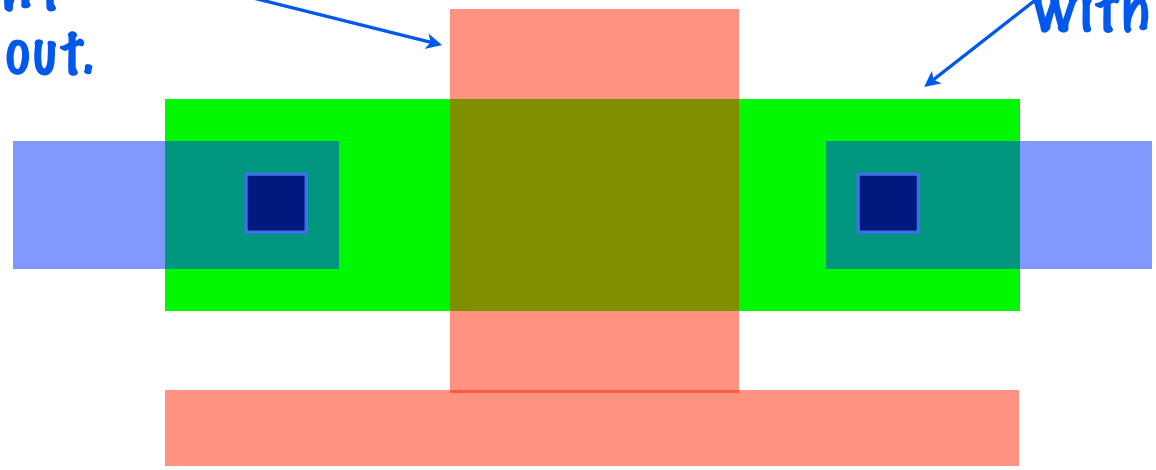
Design Rules for Masks (circa 1986)

Poly overhang. So that if masks are misaligned, channel doesn't short out.

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

length

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



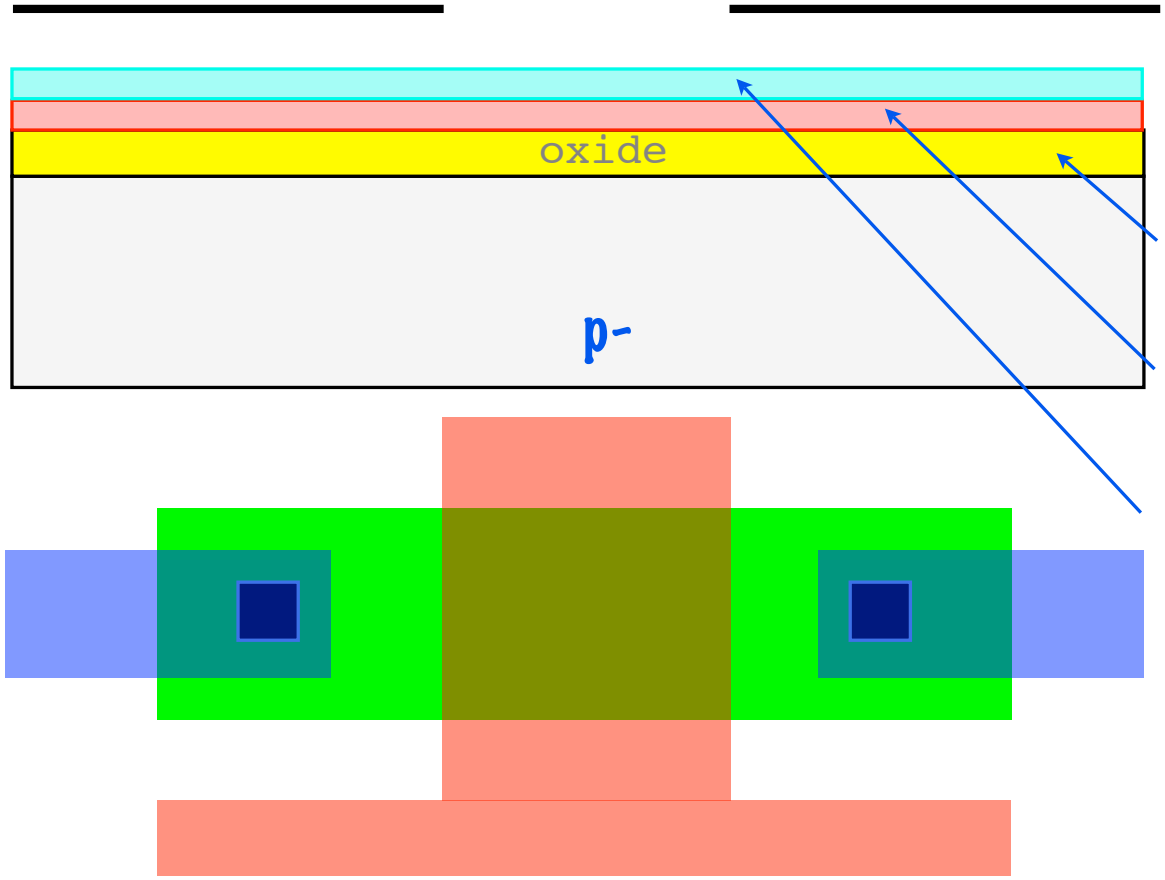
#1: n+ diffusion
#2: poly (gate)

#3: diff contact
#4: metal

Adapted from [Asanovic'11]

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 undoped polysilicon

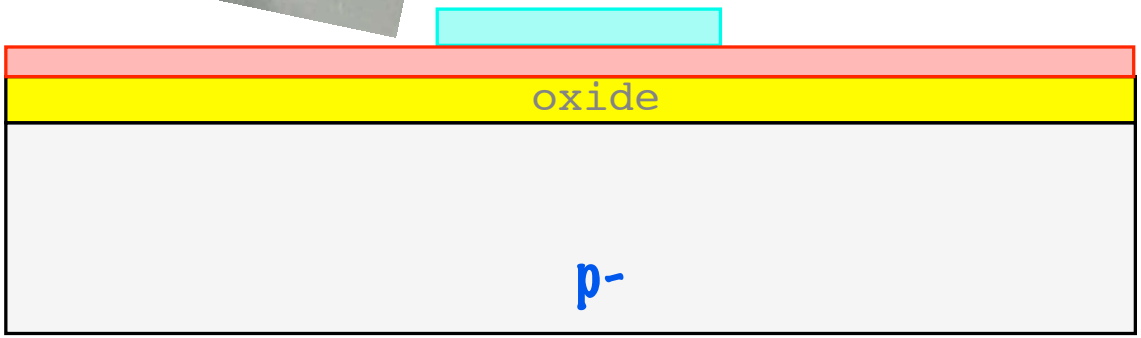
#4: spin on photoresist

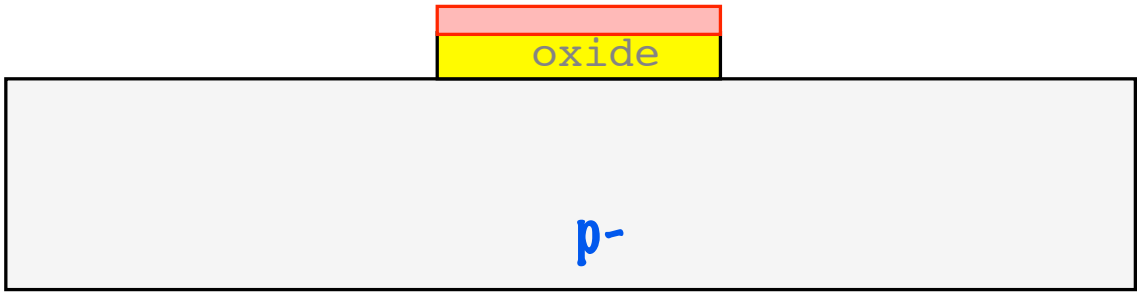
#5: place positive poly mask and expose with UV.

Adapted from [Asanovic'11]

Wet Etch to Remove Unmasked Regions

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

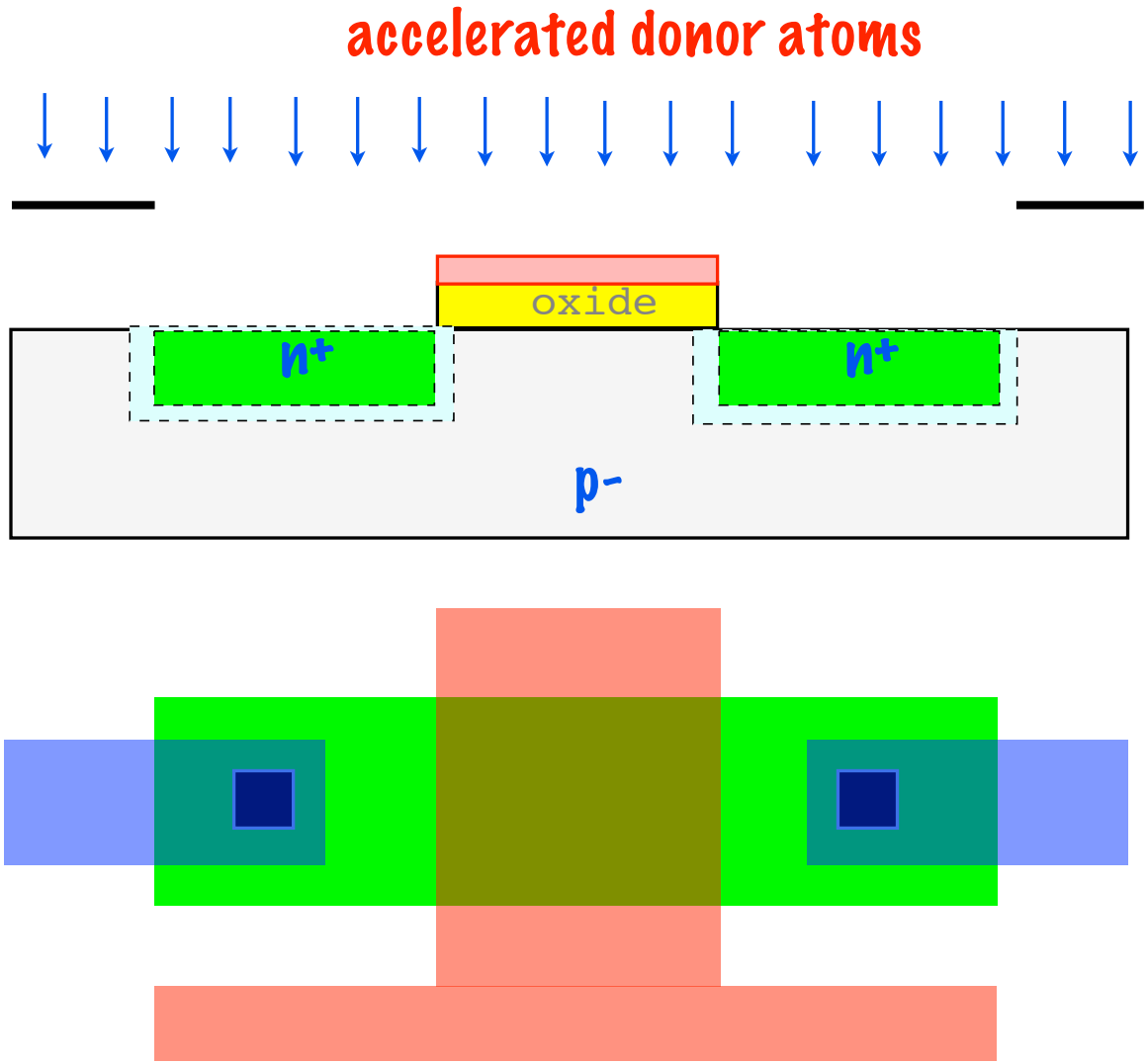


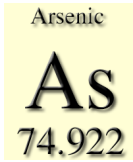


After etch and resist removal

Adapted from [Asanovic'11]

Use Diffusion Mask to Implant N-Type



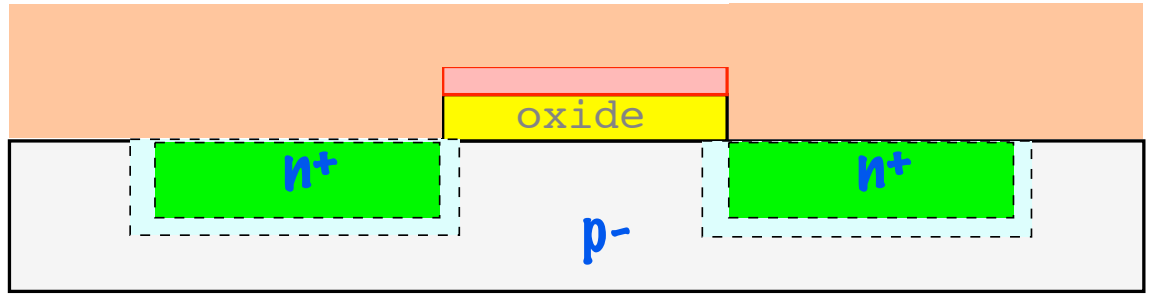


Notice how donor atoms are blocked by gate and do not enter channel.

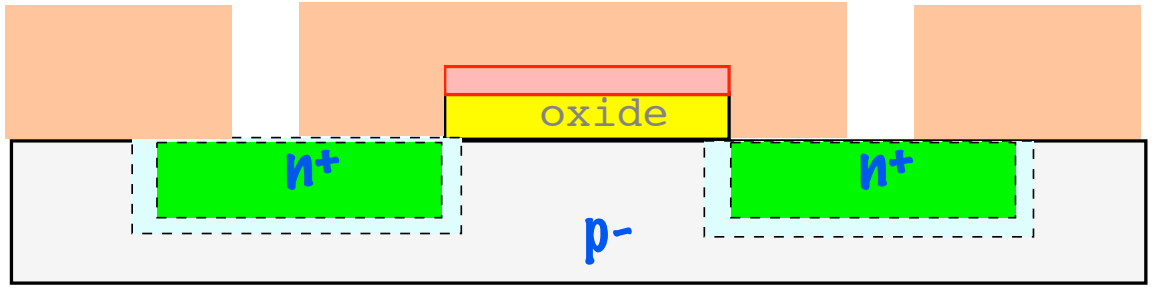
Thus, the channel is "self-aligned", precise mask alignment is not needed!

Adapted from [Asanovic'11]

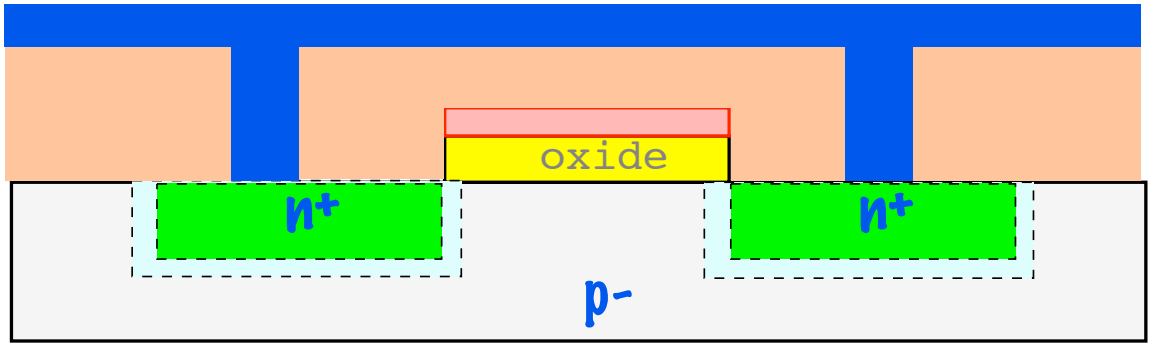
Metallization Completes Device



Grow a thick oxide on top of the wafer.



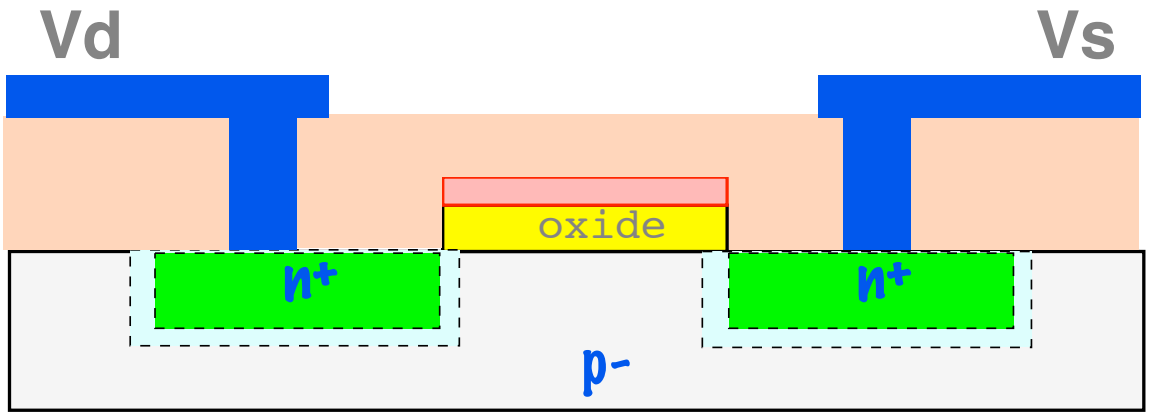
Mask and etch to make contact holes



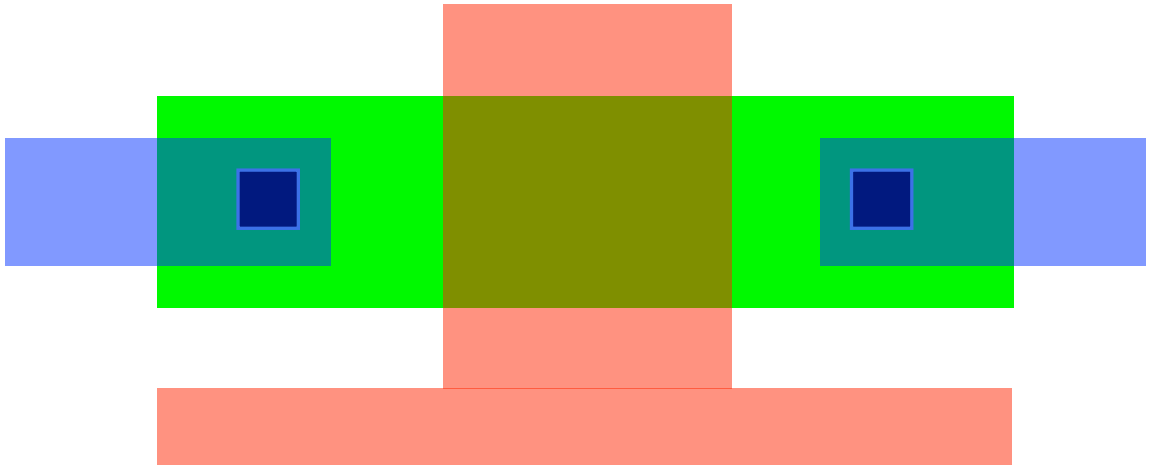
Put a layer of metal on chip. Be sure to fill in the holes!

Adapted from [Asanovic'11]

Final NMOS Transistor



Top-down view:

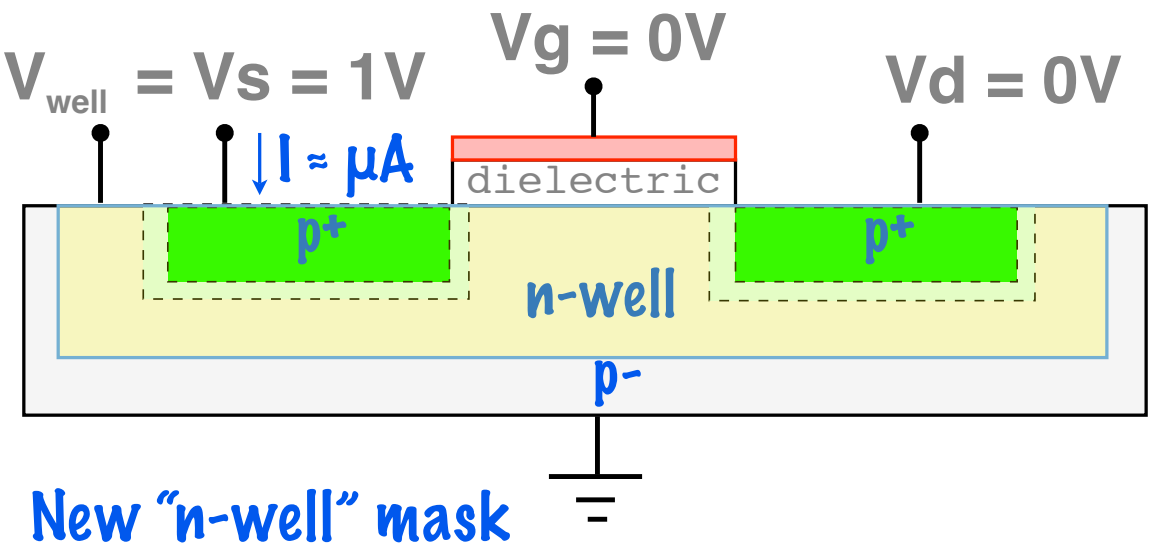


"The planar process"

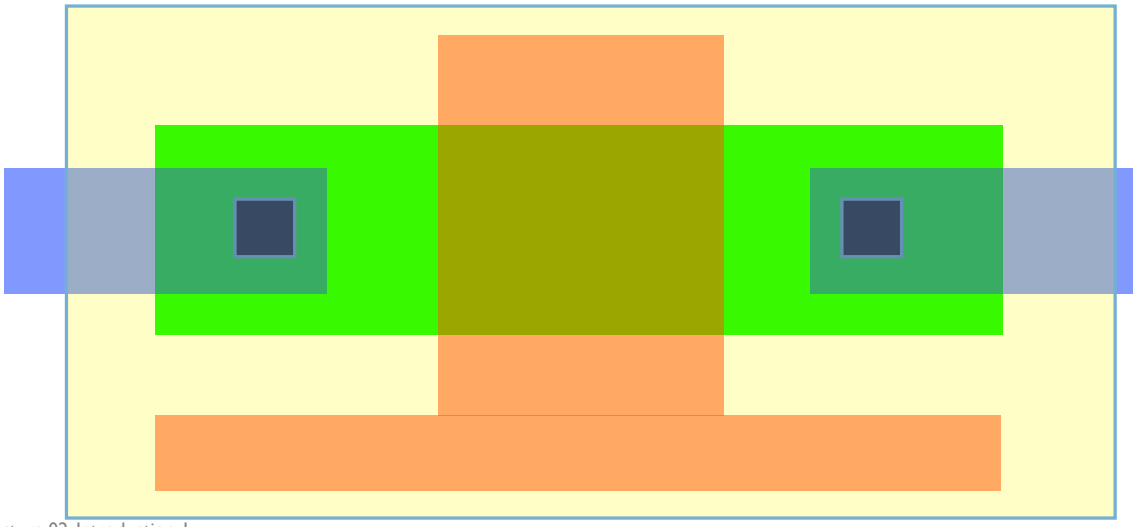
Jean Hoerni, Fairchild Semiconductor 1958

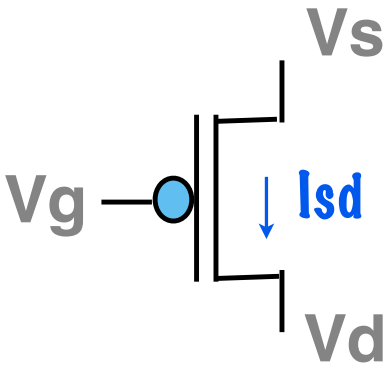
Adapted from [Asanovic'11]

PMOS Transistor is Dual of NMOS Transistor



New "n-well" mask



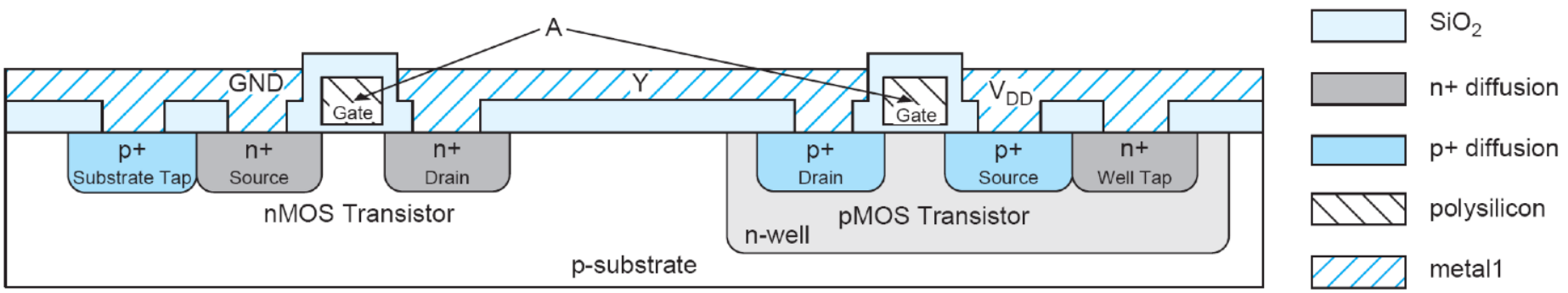


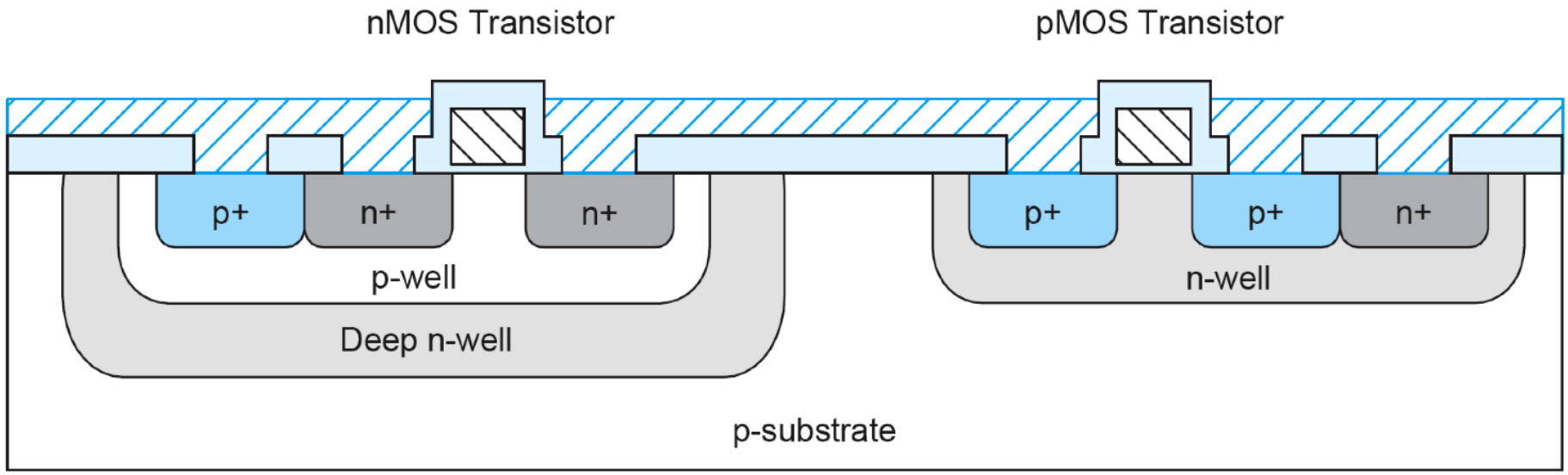
"Mobility" of holes is slower than electrons.

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

Adapted from [Asanovic'11]

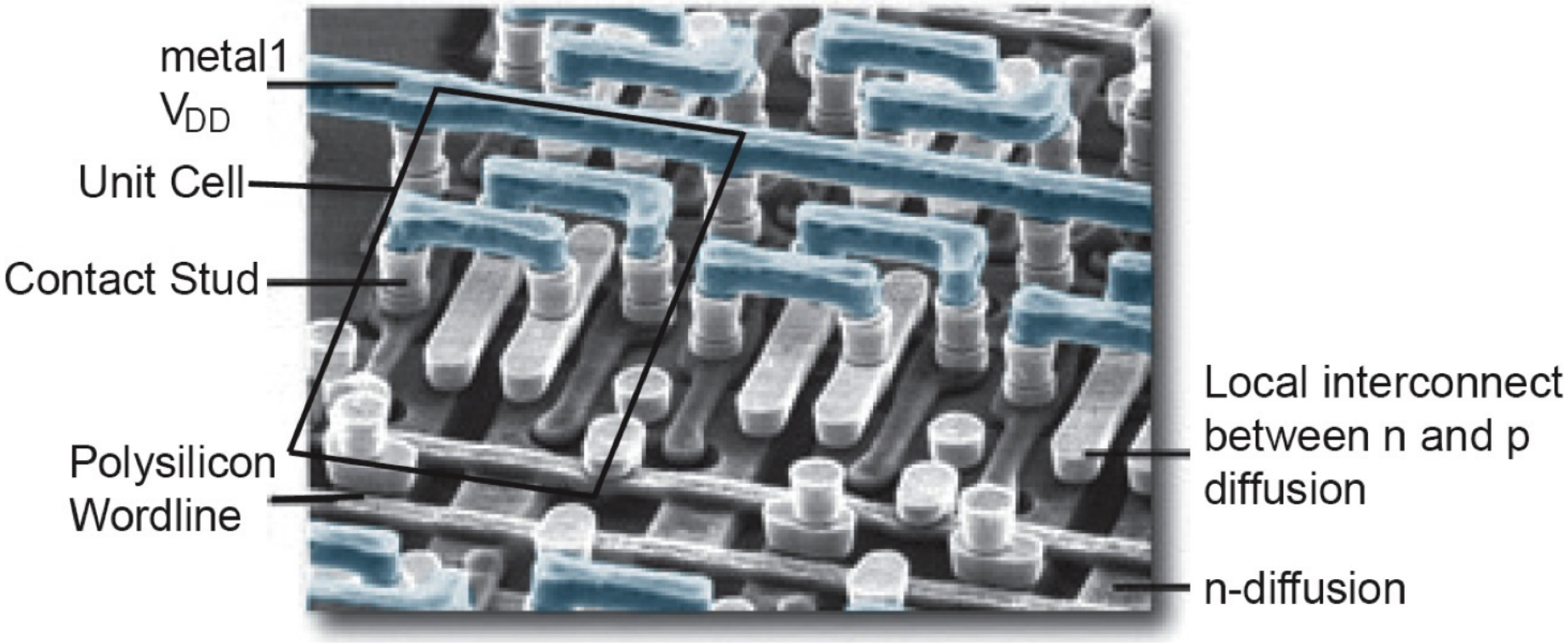
Single- and Triple-Well Processes





Adapted from [Weste'11]

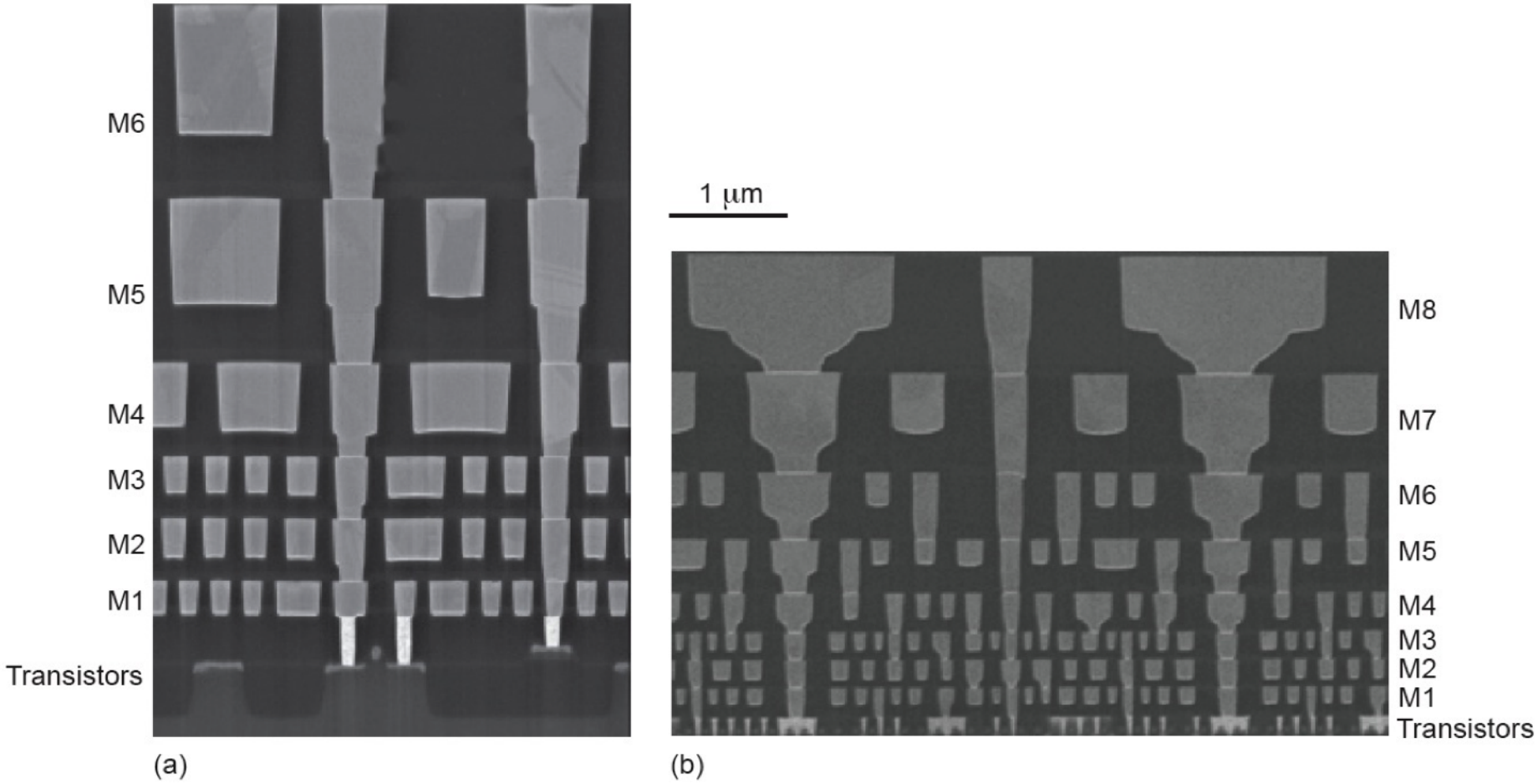
Local Interconnect



IBM 6-Transistor SRAM Cell

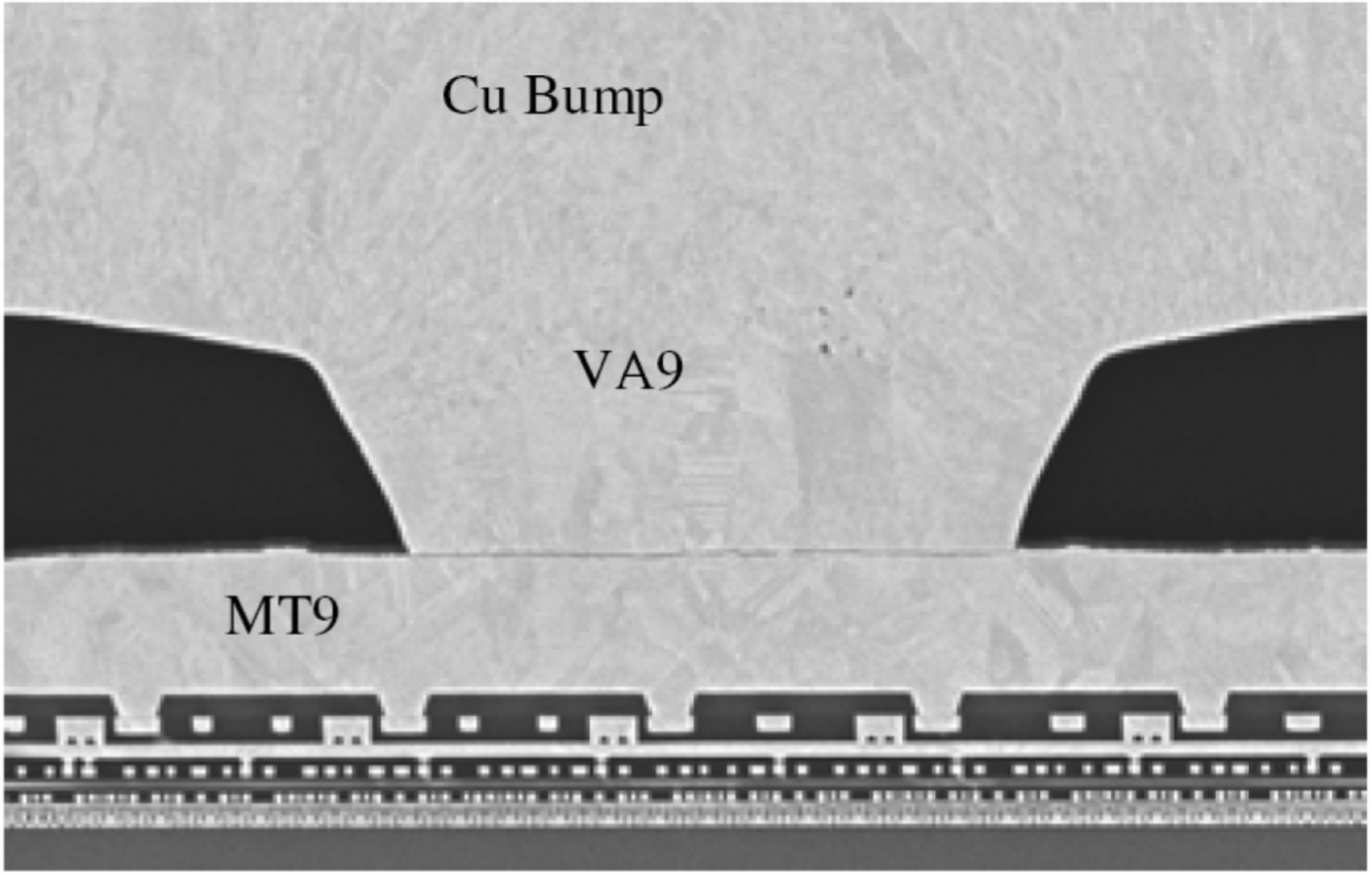
Adapted from [Weste'11]

Intel Metal Stacks: 90nm and 45nm



Adapted from [Weste'11]

Intel Metal Stacks: 45nm with M9 and I/O Bump



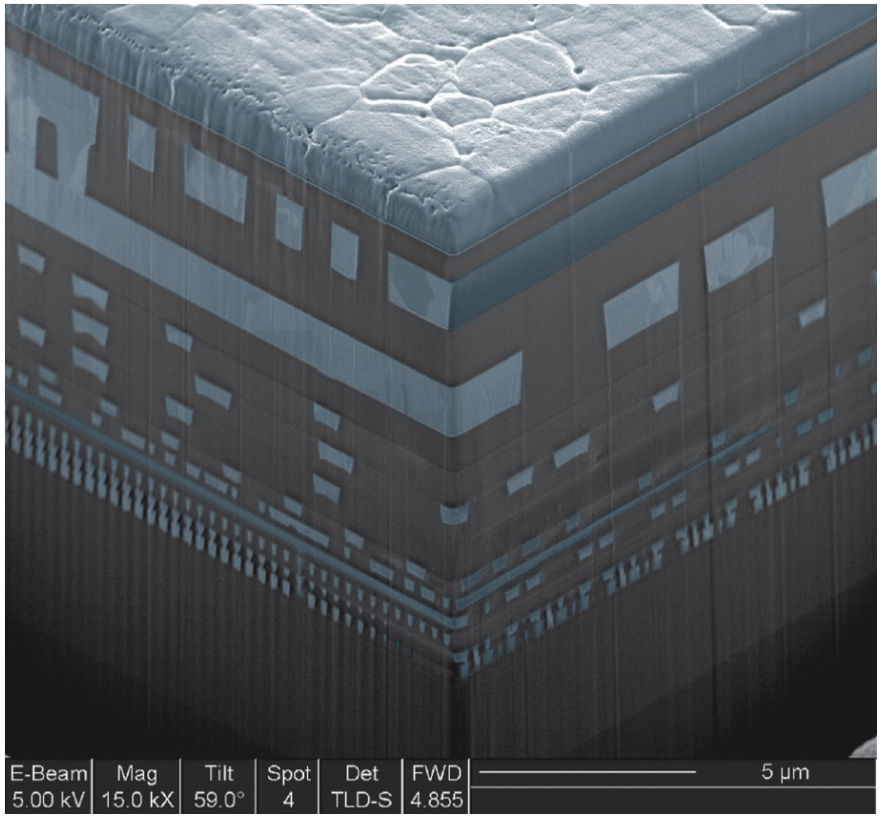
Adapted from [Weste'11]

Intel Metal Layer Dimensions in 45nm

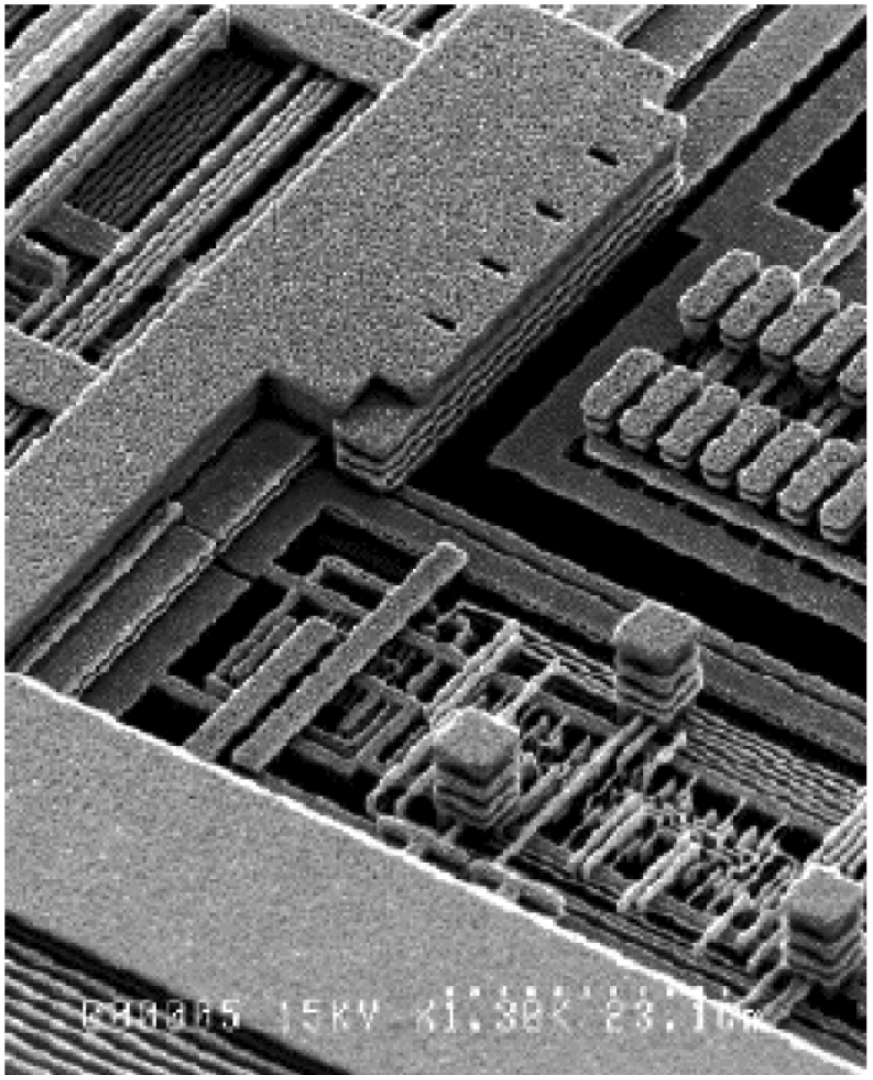
Layer	t (nm)	w (nm)	s (nm)	pitch (nm)
M9	7µm	17.5µm	13µm	30.5µm
M8	720	400	410	810
M7	504	280	280	560
M6	324	180	180	360
M5	252	140	140	280
M4	216	120	120	240
M3	144	80	80	160
M2	144	80	80	160
M1	144	80	80	160

Adapted from [Weste'11]

IBM Metal Stacks



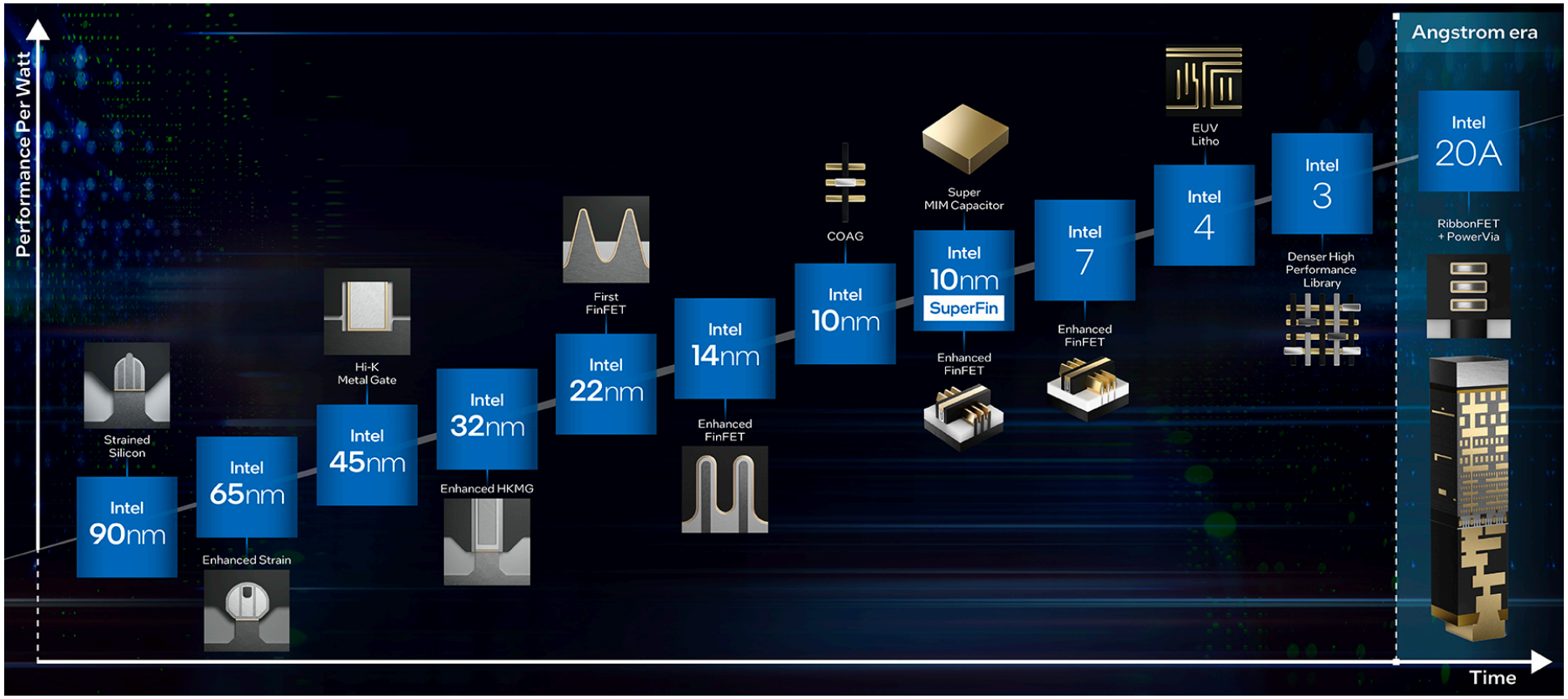
IBM 11-layer Copper Metal Stack



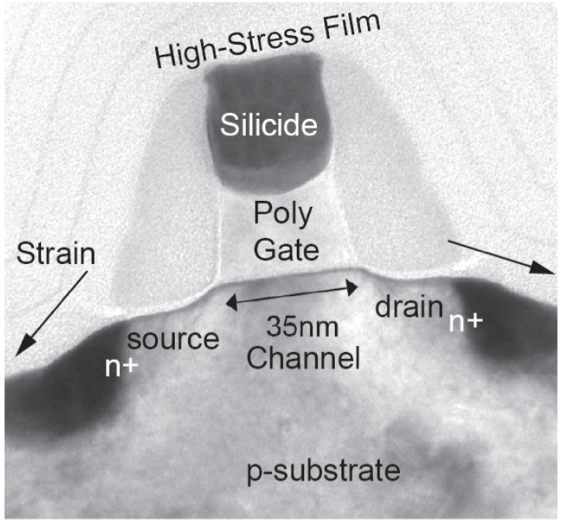
IBM 6-layer Copper Metal Stack

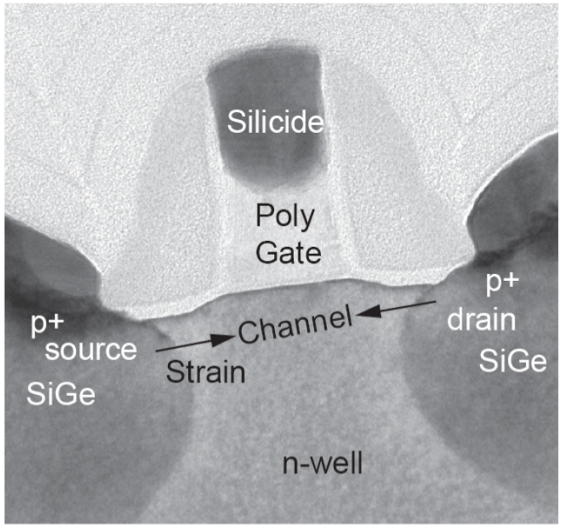
Adapted from [Weste'11]

Technology Scaling via Process Enhancements



Process Enhancements

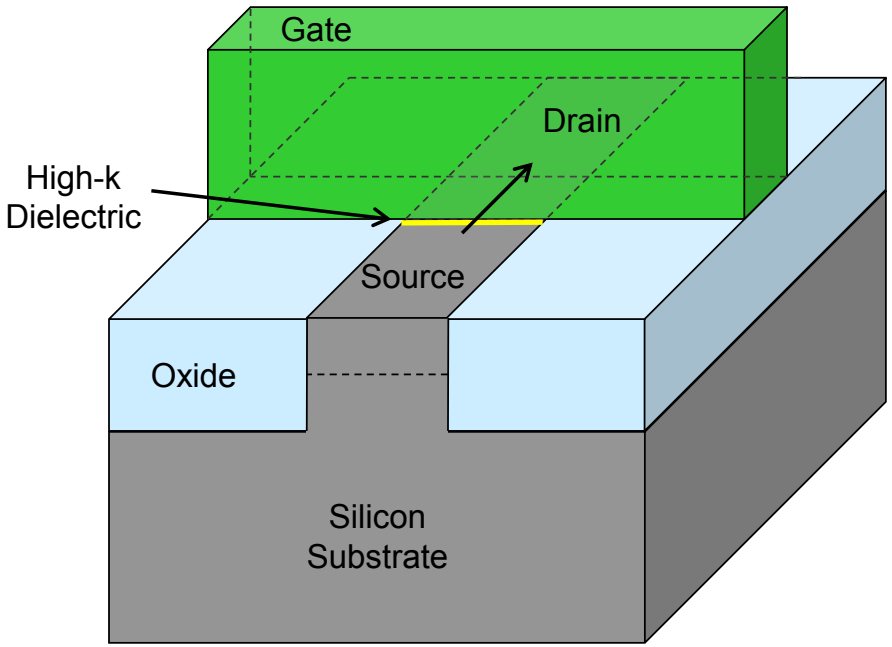


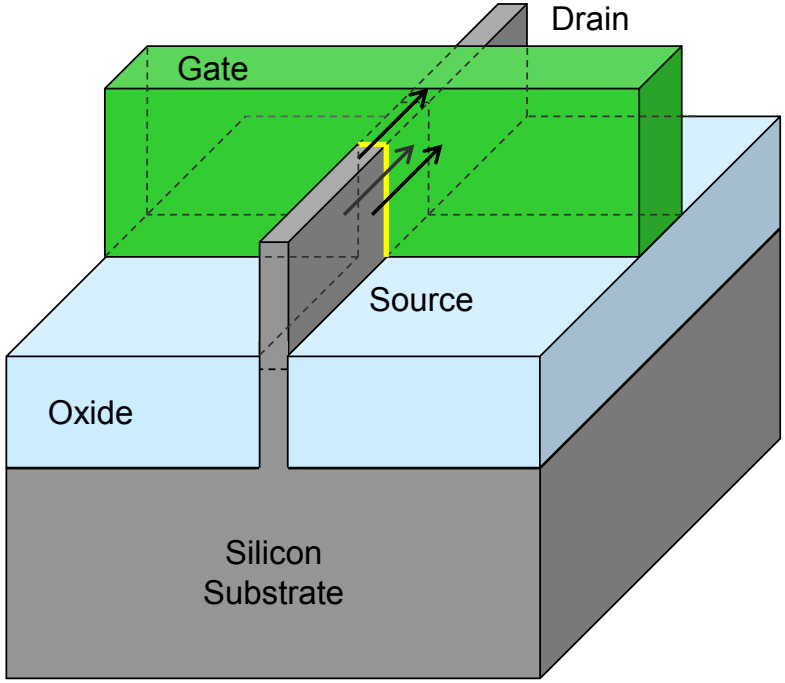


- **High-K Dieletrics and Metal Gates –** Replacing silicon dioxide gate dielectric with a high-K material allows increased vertical electric field without increasing gate leakage
- **Strained Silicon –** Layer of silicon in which silicon atoms are stretched beyond their normal interatomic distance leading to better mobility
- **Gate Engineering –** Multiple transistor designs with different threshold voltages to allow optimization of delay or power

Adapted from [Asanovic'11,Weste'11]

FinFET Transistors



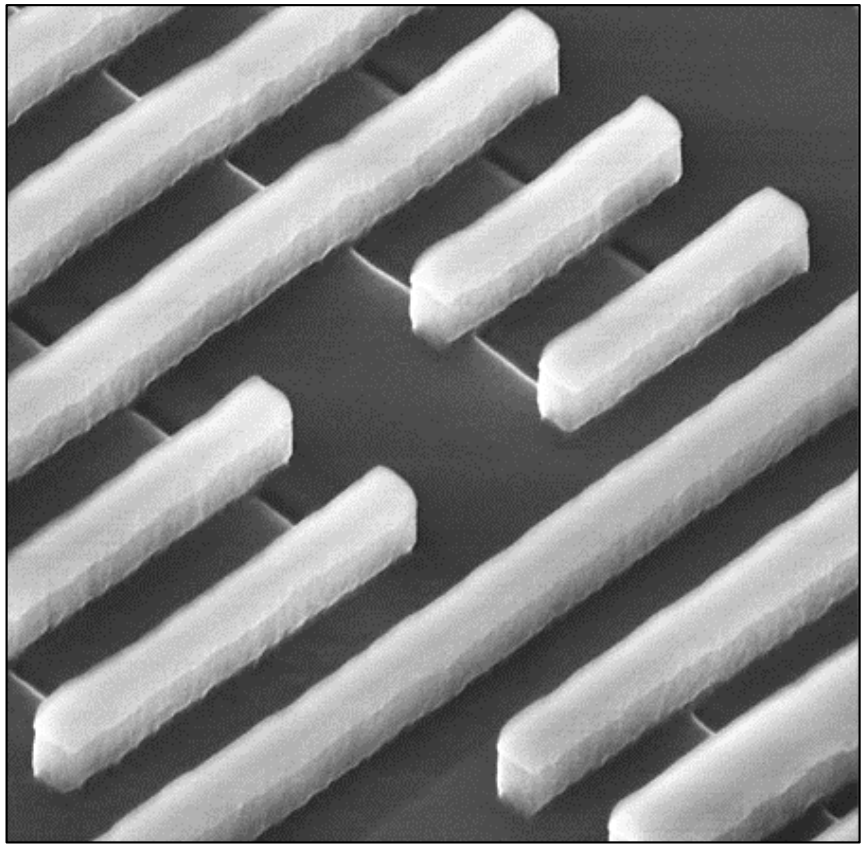


2D planar transistors form a conducting channel in silicon region under the gate electrode

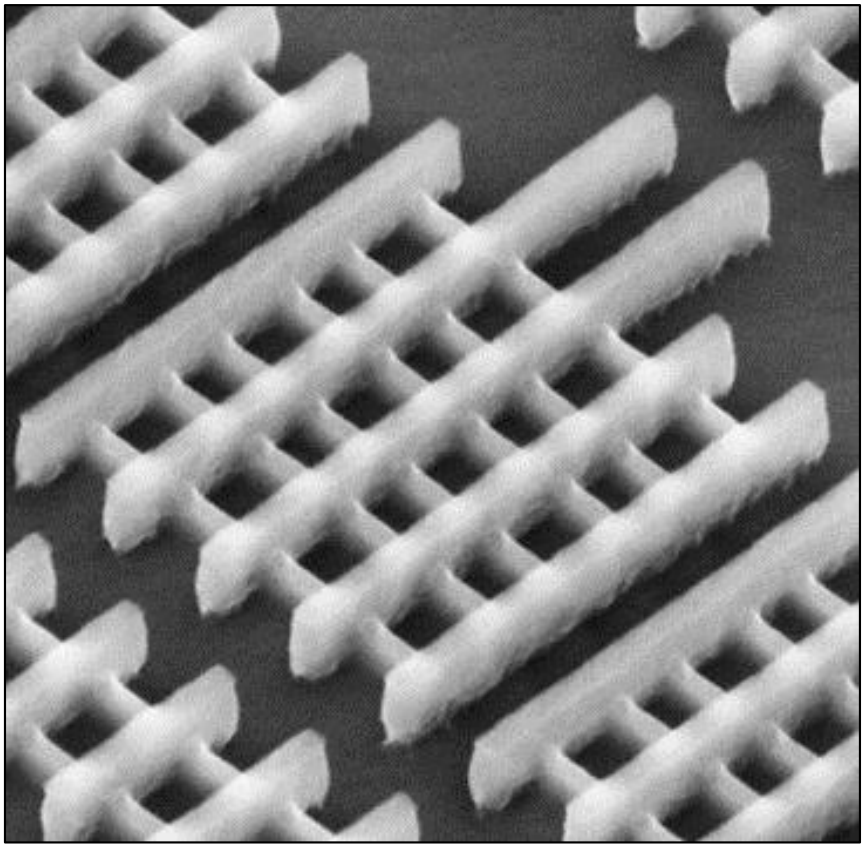
3D tri-gate transistors form conducting channels on three sides of a vertical fin structure

FinFET Transistors

32 nm Planar Transistors

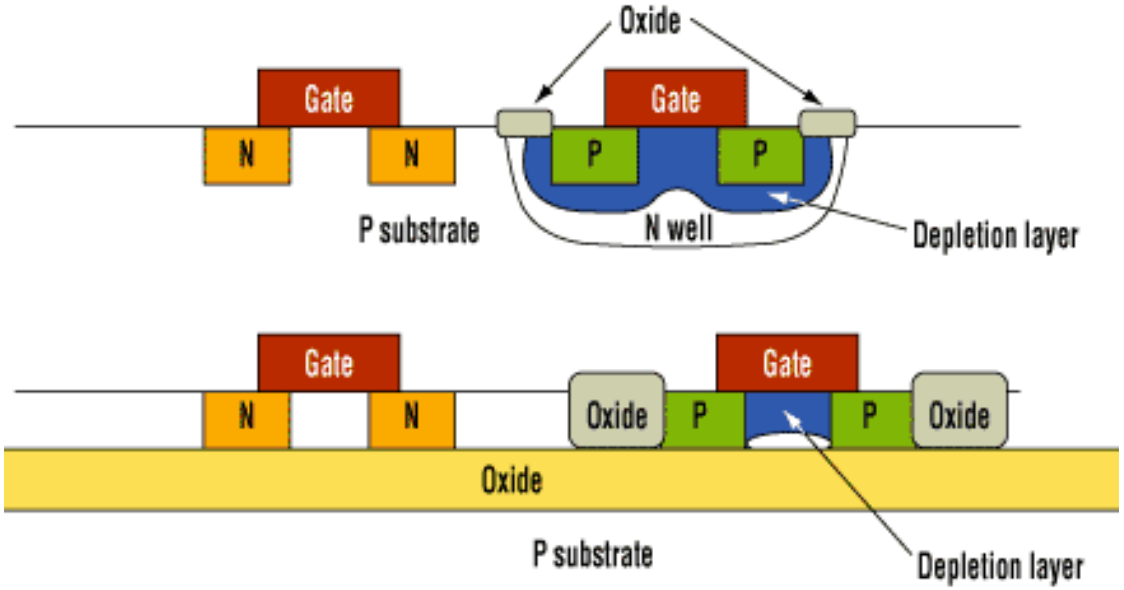


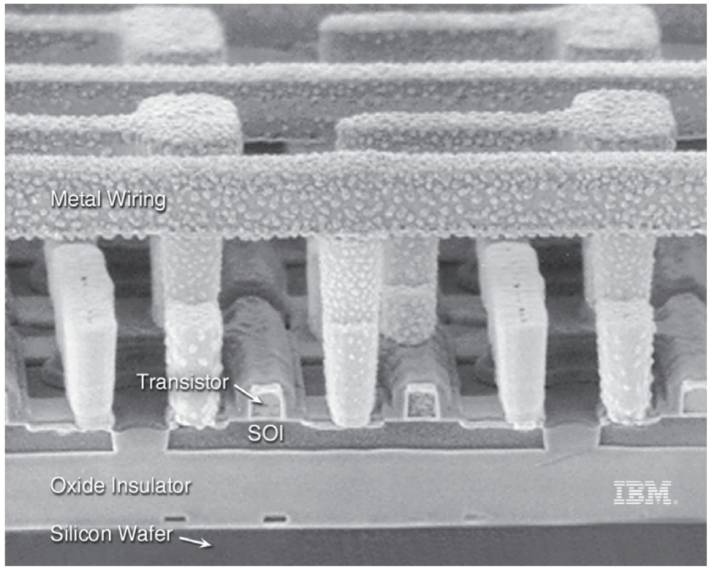
22 nm Tri-Gate Transistors



Bulk vs. Silicon-on-Insulator Processing

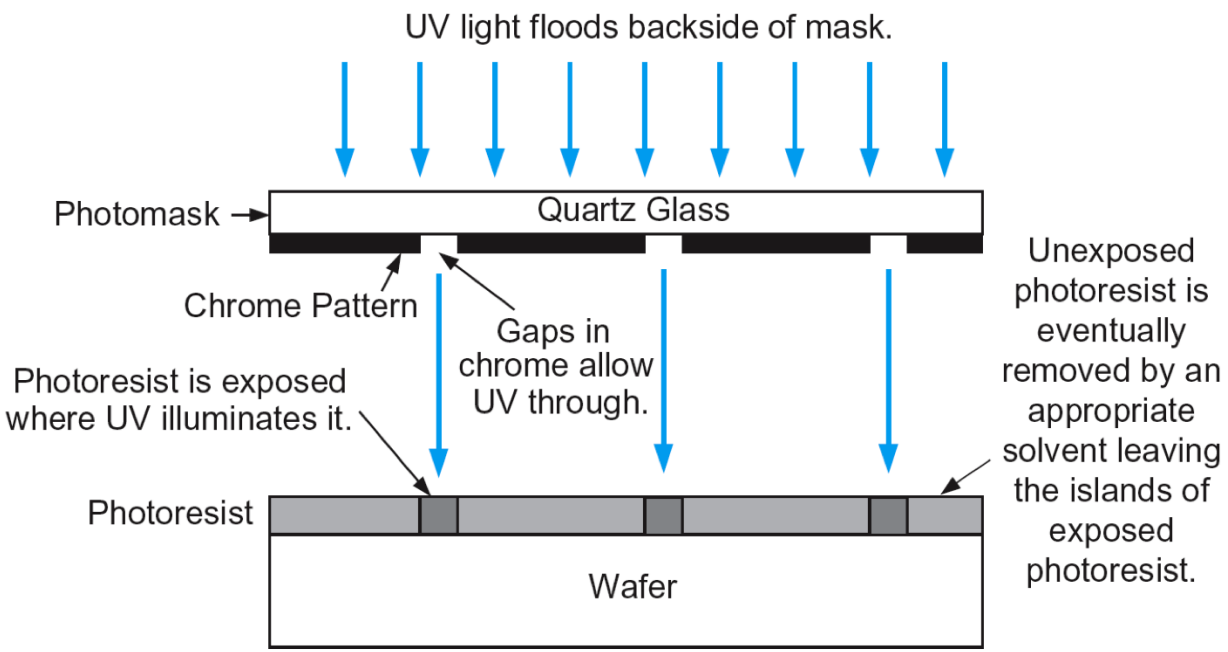
- Eliminates parasitic capacitance between source/drain and the body → lower energy, higher performance
- Lower subthreshold leakage, but threshold voltage varies over time
- 10–15% increase in total manufacturing cost due to substrate cost

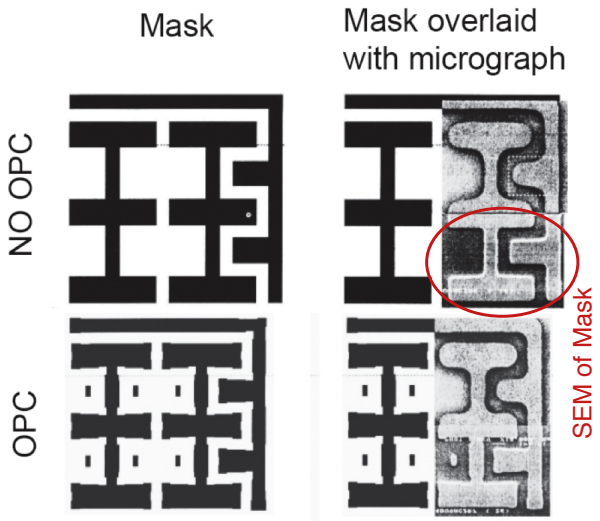


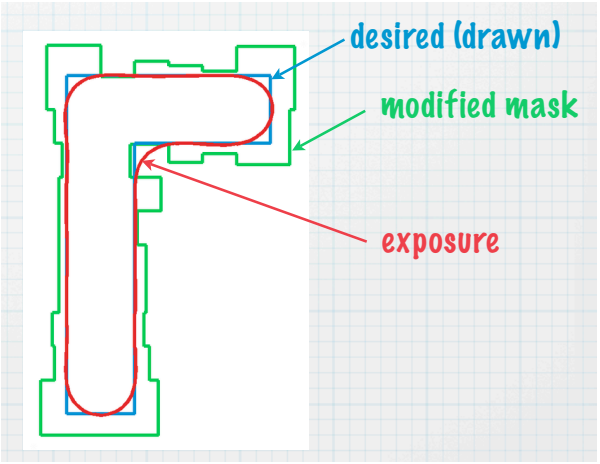


Adapted from [Asanovic'11,Weste'11]

Lithography



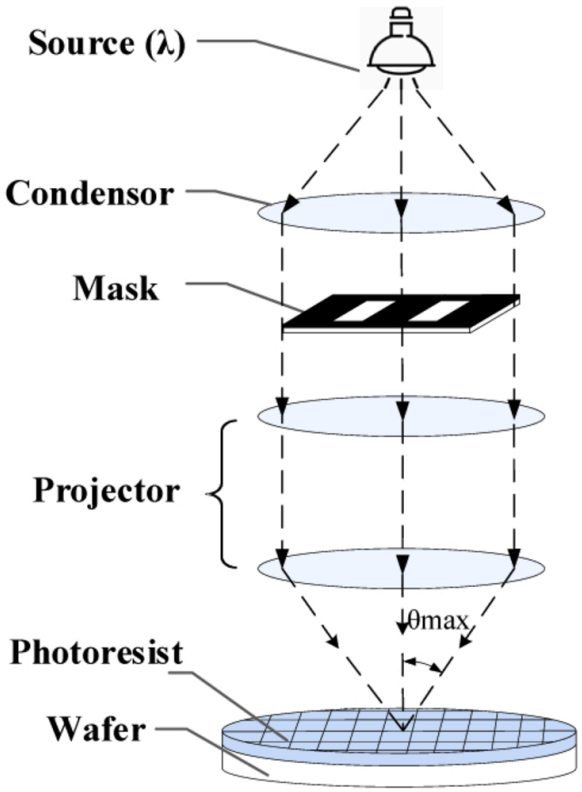


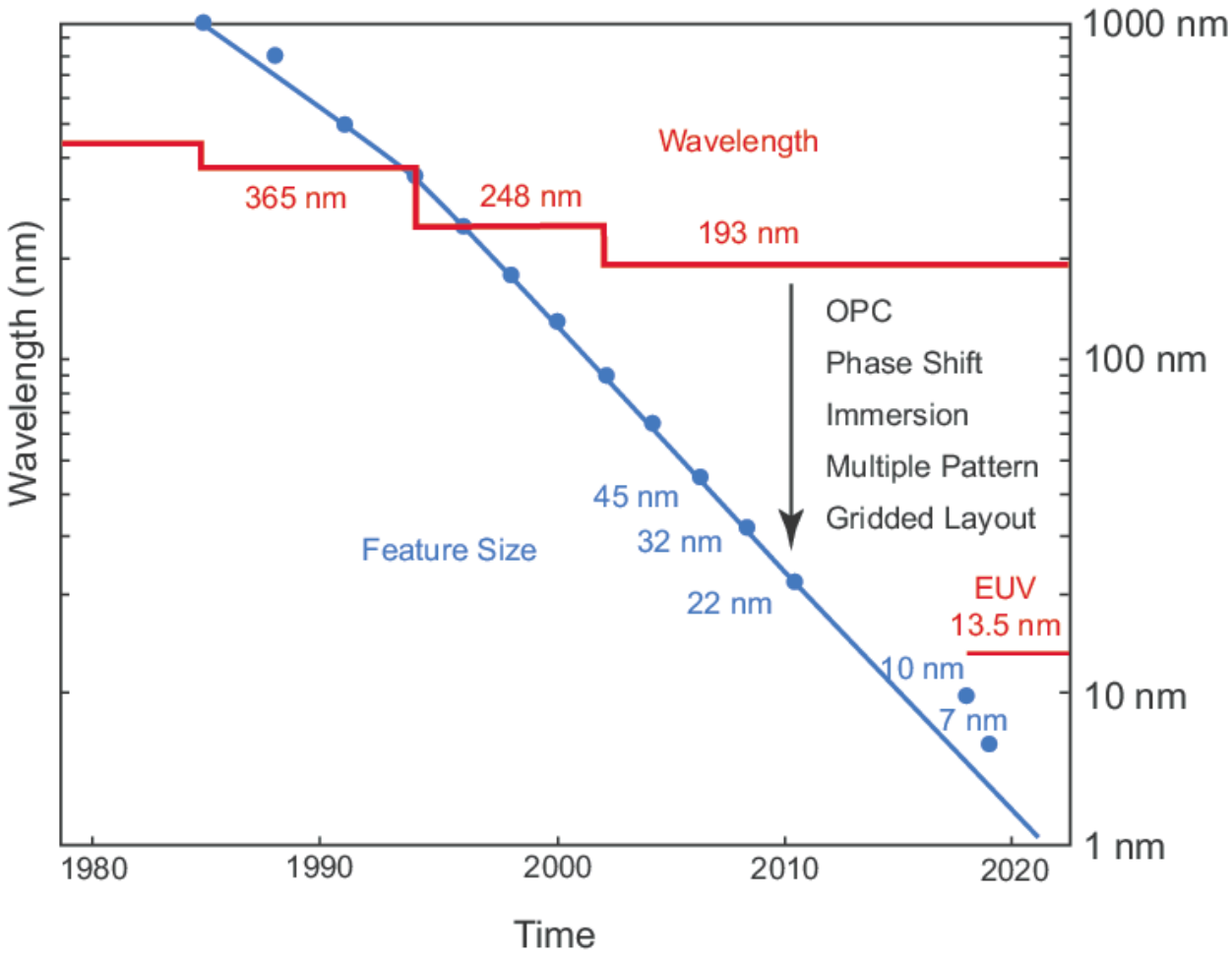


- Resolution of patterns far exceeds wavelength of light used for exposure which is usually 193 nm generated with an argon fluoride laser
- Sophisticated tricks used to pattern 10–100 µm features including immersion lithography, optical proximity correction, double patterning

Adapted from [Asanovic'11,Weste'11]

Extreme Ultraviolet Lithography





Take-Away Points

- Although a basic understanding of devices and fabrication is important for understanding technology constraints, mostly in this course we will focus on first-order RC models of CMOS logic, state, and interconnect
- In the next topic of this part of the course, we will briefly introduce CMOS circuits using these devices
  - Combinational Logic: static CMOS, pass-transistor, tri-state buffers
  - Sequential State: latches, flip-flops
- In the next part of the course, we will explore the details of how to quantitatively evaluate the cycle time, area, and energy of these digital circuits

Acknowledgments

- [Weste'11] N. Weste and D. Harris, "CMOS VLSI Design: A Circuits and Systems Perspective," 4th ed, Addison Wesley, 2011.
- [Terman'02] C. Terman and K. Asanović, "CMOS Technology," and "Wires, " MIT 6.371 Introduction to VLSI Systems, Lectures, 2002.
- [Asanovic'11] K. Asanović, J. Wawrzynek, and J. Lazzaro, "Introduction," UC Berkeley CS 250 VLSI Systems Design, Lecture, 2011.