

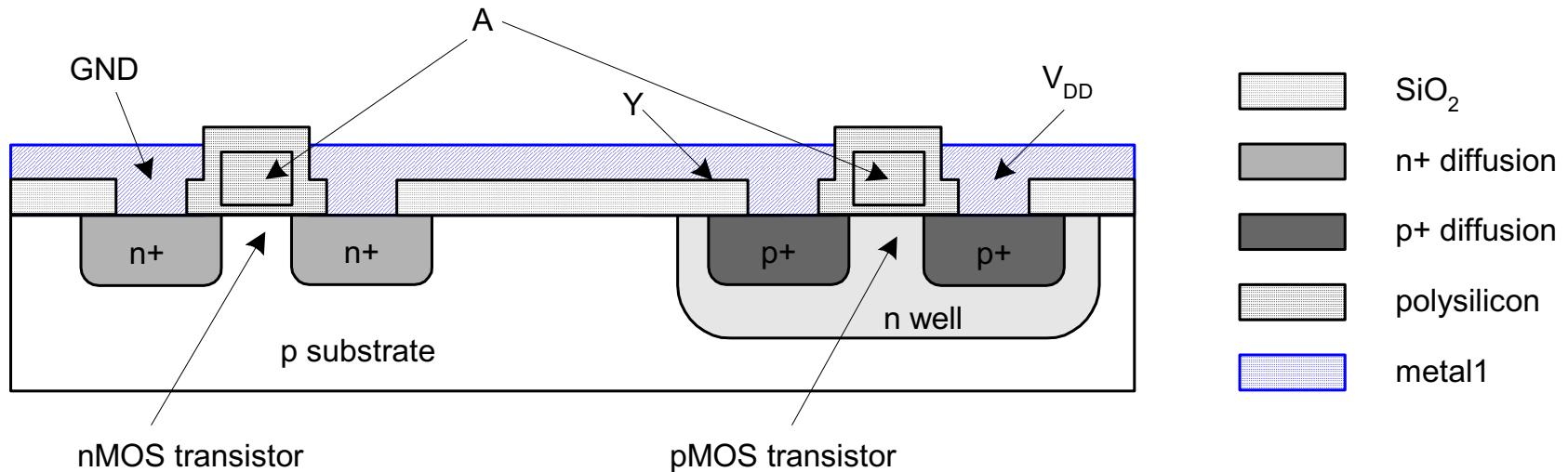
# **ECE1388 VLSI Design Methodology**

## **Lecture 2: CMOS Fabrication & Layout**

Adapted from Weste, Harris CMOS VLSI Design

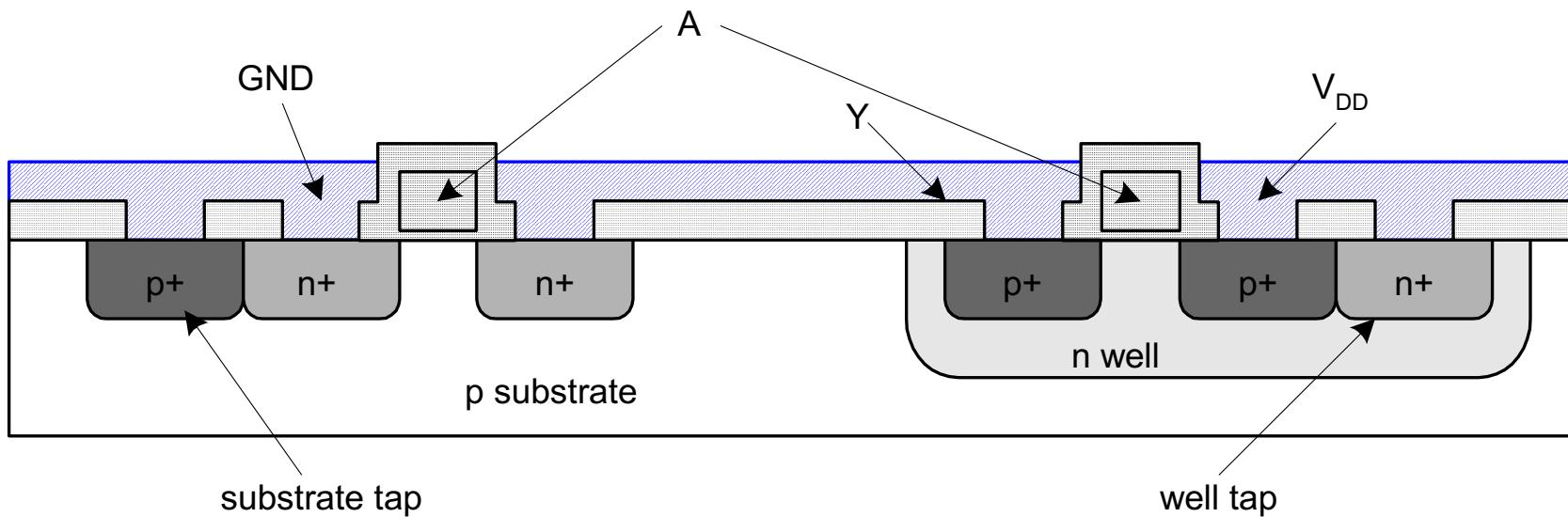
# Inverter Cross-section

- Typically use p-type substrate for nMOS transistors
- Requires n-well for body of pMOS transistors



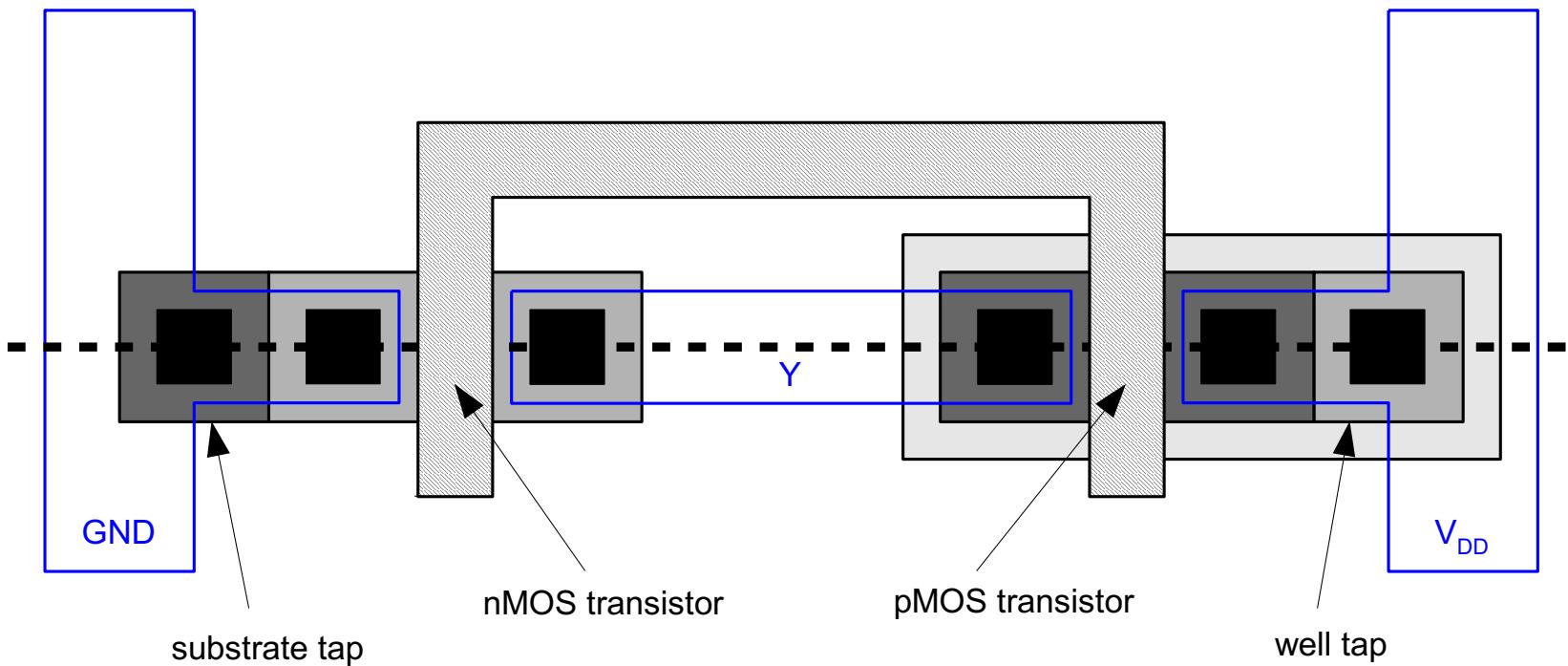
# Well and Substrate Taps

- Substrate must be tied to GND and n-well to  $V_{DD}$
- Metal to lightly-doped semiconductor forms poor connection called Shottky Diode
- Use heavily doped well and substrate contacts / taps



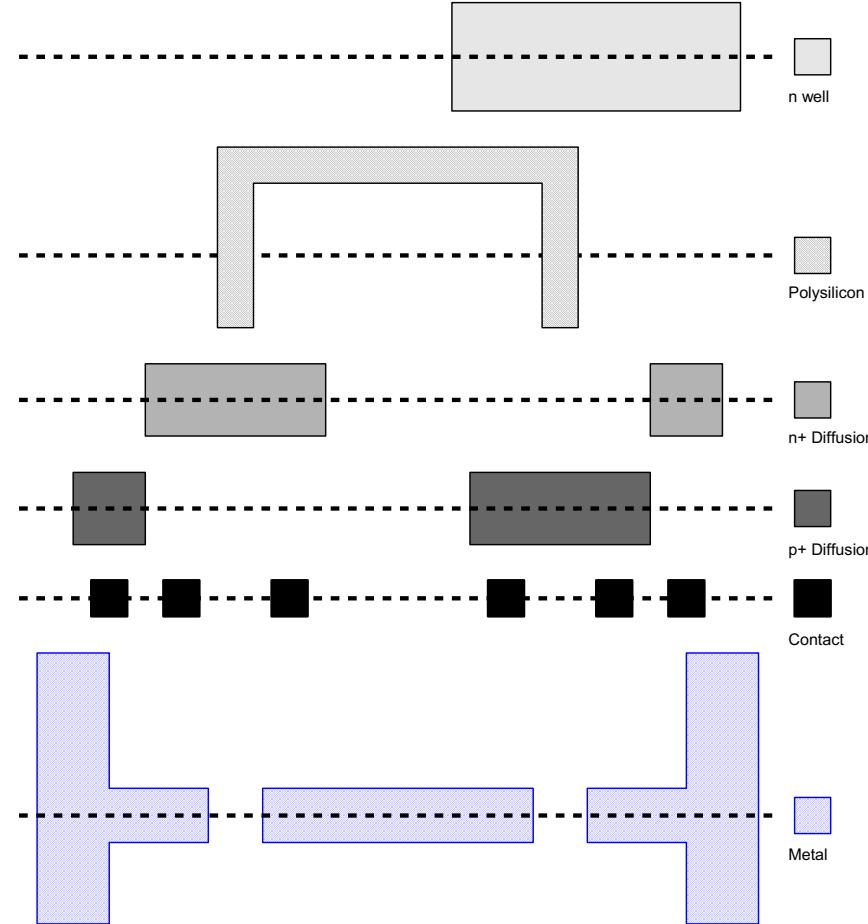
# Inverter Mask Set

- Transistors and wires are defined by *masks*
- Cross-section taken along dashed line



# Detailed Mask Views

- n-well
- Polysilicon
- n+ diffusion
- p+ diffusion
- Contact
- Metal



# Fabrication Steps

- Start with blank wafer
- Build inverter from the bottom up
- First step will be to form the **n-well**
  - Cover wafer with protective layer of  $\text{SiO}_2$  (oxide)
  - Remove layer where n-well should be built
  - Implant or diffuse n dopants into exposed wafer
  - Strip off  $\text{SiO}_2$



p substrate

# Oxidation

- Grow  $\text{SiO}_2$  on top of Si wafer
  - auxiliary layer for substrate protection
  - will be removed
  - will not be included in MOSFET structure
  - 900 – 1200 C with  $\text{H}_2\text{O}$  or  $\text{O}_2$  in oxidation furnace



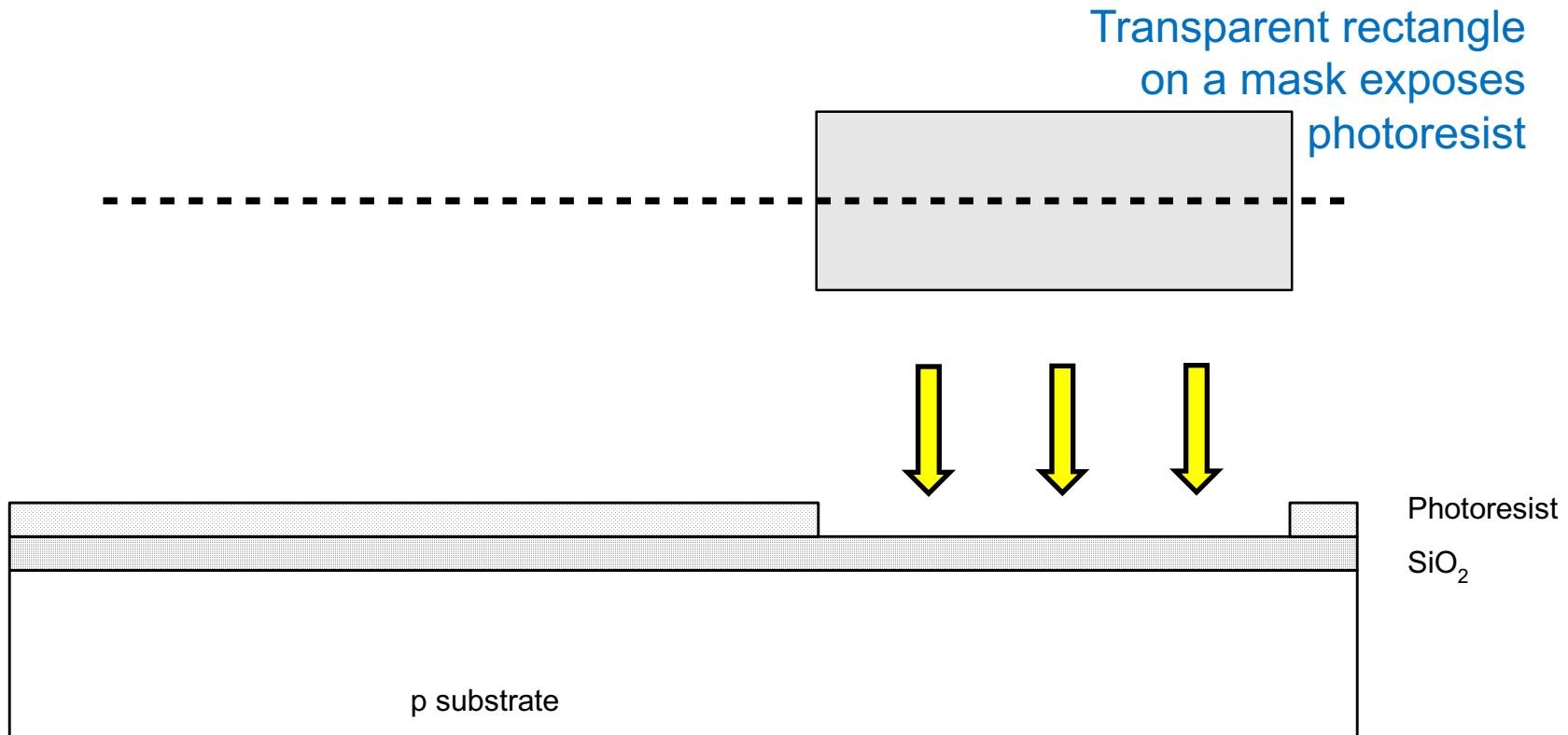
# Photoresist

- Spin on photoresist
  - Photoresist is a light-sensitive organic polymer
  - Softens where exposed to light



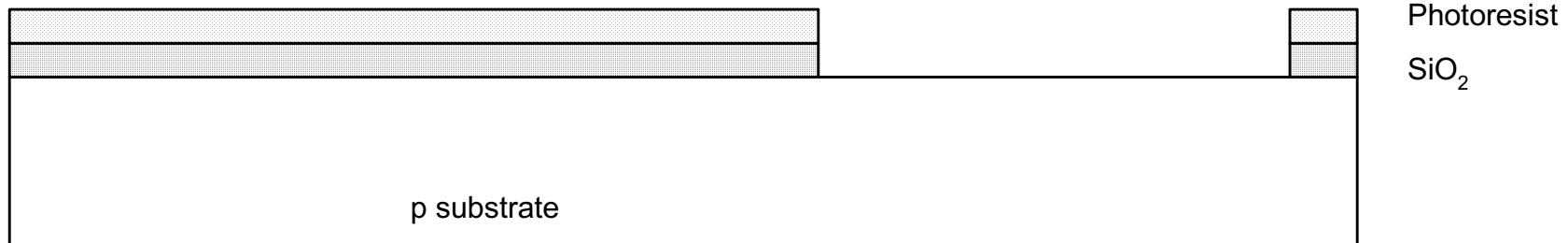
# Lithography

- ❑ Expose photoresist through n-well mask
- ❑ Strip off exposed softened photoresist



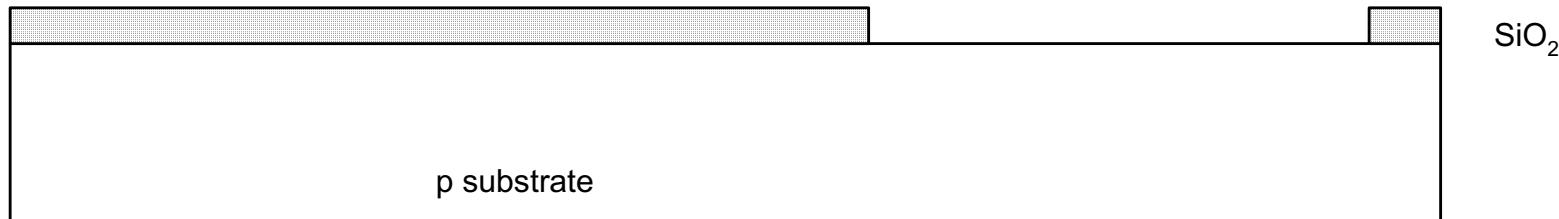
# Etch

- Etch oxide with hydrofluoric acid (HF)
  - Seeps through skin and eats bone; nasty stuff!!!
- Only attacks oxide where photoresist has been removed



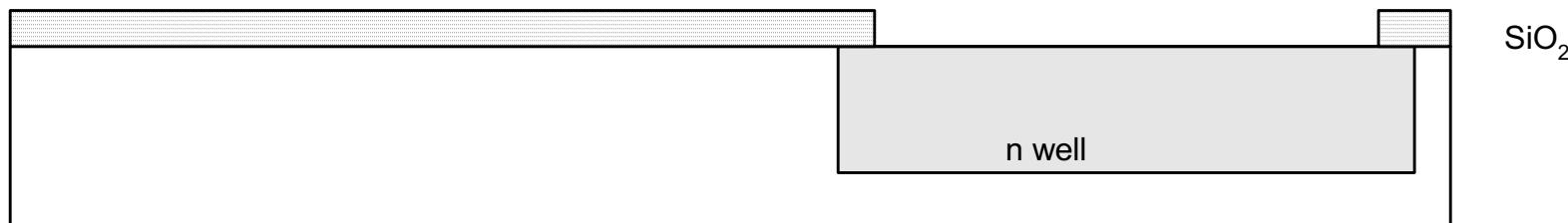
# Strip Photoresist

- Strip off remaining photoresist
  - Use mixture of acids called piranah etch
- Necessary so resist doesn't melt in next step



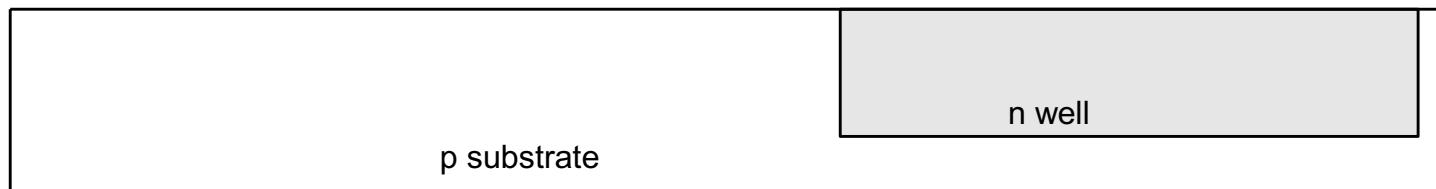
# n-well

- ❑ n-well is formed with diffusion or **ion implantation**
- ❑ Diffusion
  - Place wafer in furnace with arsenic gas
  - Heat until As atoms diffuse into exposed Si
- ❑ Ion Implantation
  - Blast wafer with **beam of As ions**
  - Ions blocked by  $\text{SiO}_2$ , only enter exposed Si



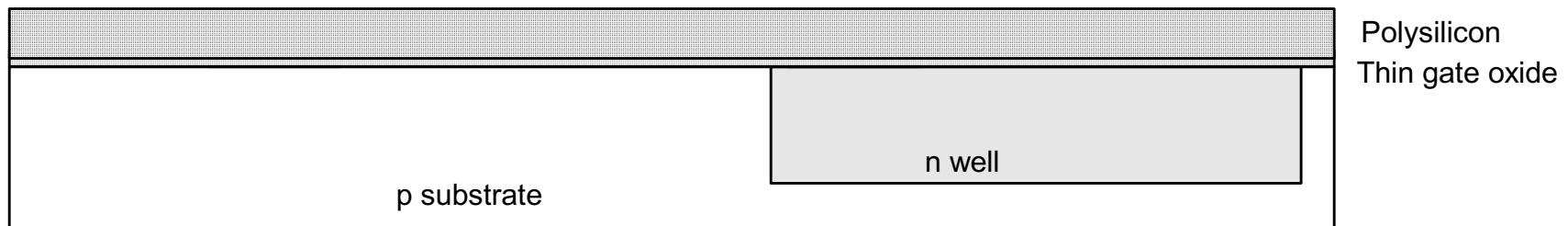
# Strip Oxide

- Strip off the remaining oxide using **hydrofluoric acid (HF)**
- Back to bare wafer with n-well
- Subsequent steps involve similar series of steps



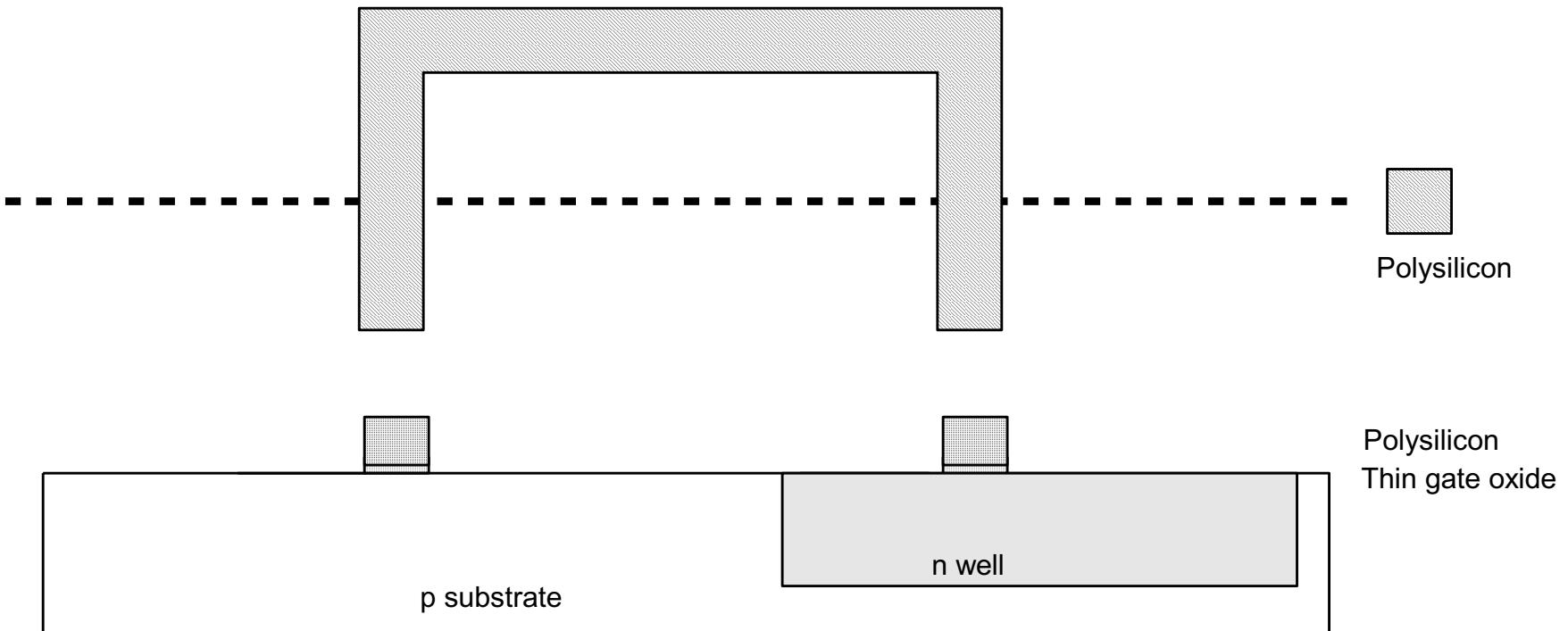
# Polysilicon

- Deposit very thin layer of gate oxide
  - < 20 Å (6-7 atomic layers)
- Chemical Vapor Deposition (CVD) of silicon layer
  - Place wafer in furnace with **Silane gas** ( $\text{SiH}_4$ )
  - Forms many small crystals called polysilicon
  - Heavily doped to be good conductor



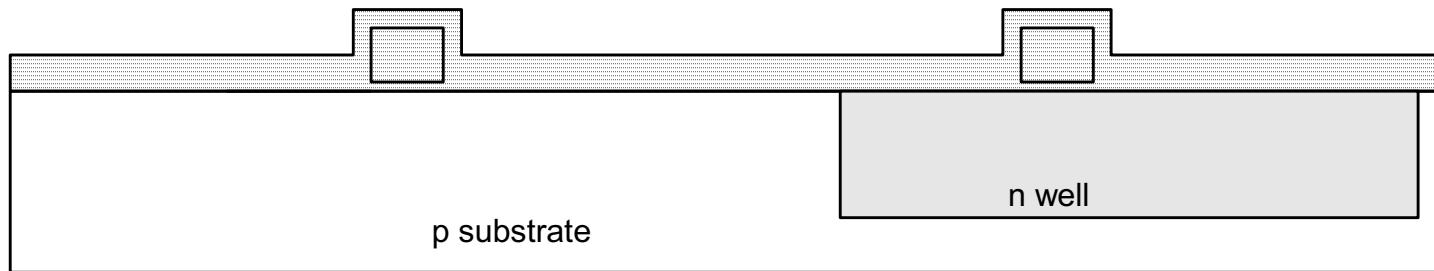
# Polysilicon Patterning

- Use same lithography process to pattern polysilicon



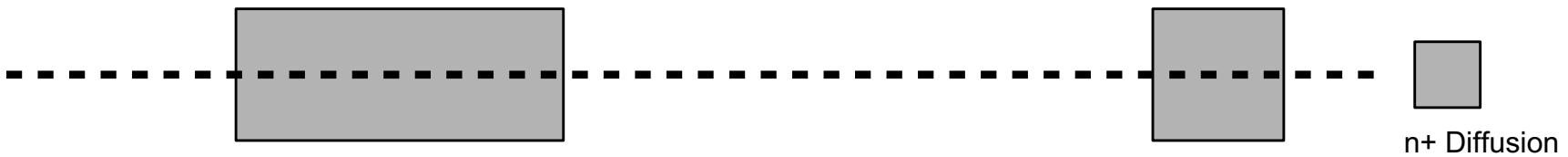
# Self-Aligned Process

- Use **both oxide and masking** to expose where n+ dopants should be diffused or implanted
- N-diffusion forms nMOS source, drain, and n-well contact

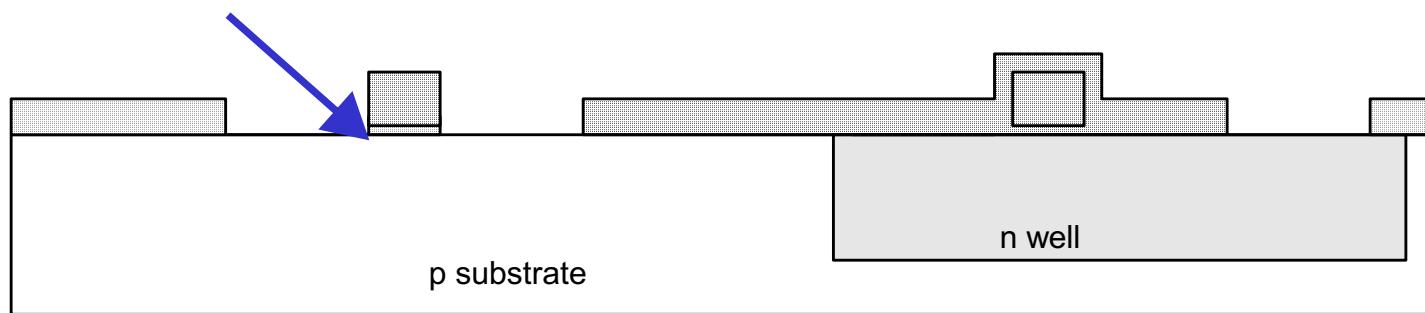


# N-diffusion

- Pattern oxide and form n+ regions
- *Self-aligned process* where gate blocks diffusion
- Polysilicon is better than metal for self-aligned gates because it does not melt during later processing

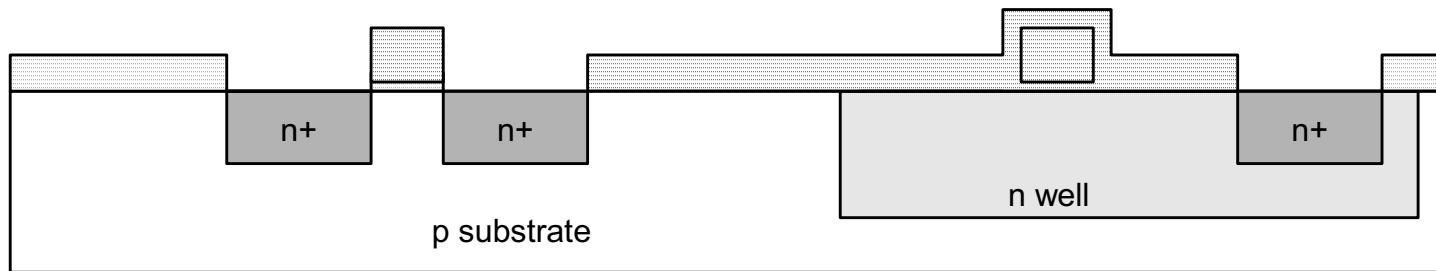


Self-aligned: gate-to-diffusion edge  
defined by gate edge, not mask



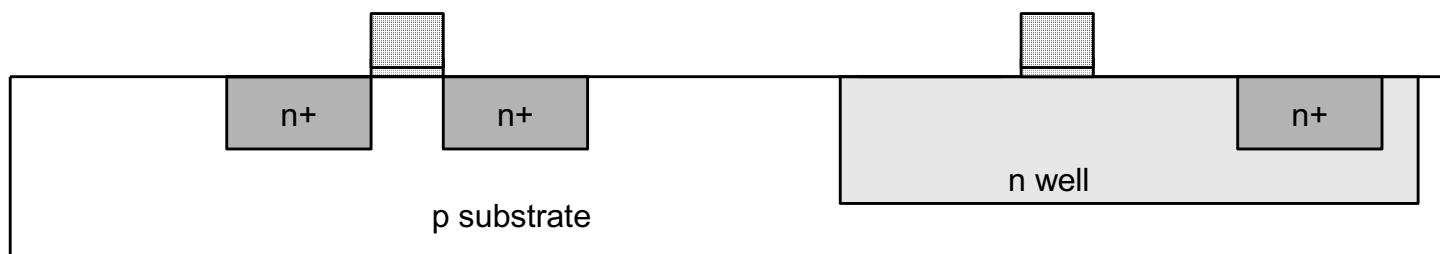
# N-diffusion cont.

- ❑ Historically dopants were diffused
- ❑ Usually ion (e.g., Phosphorus) implantation today
- ❑ But regions are still called diffusion



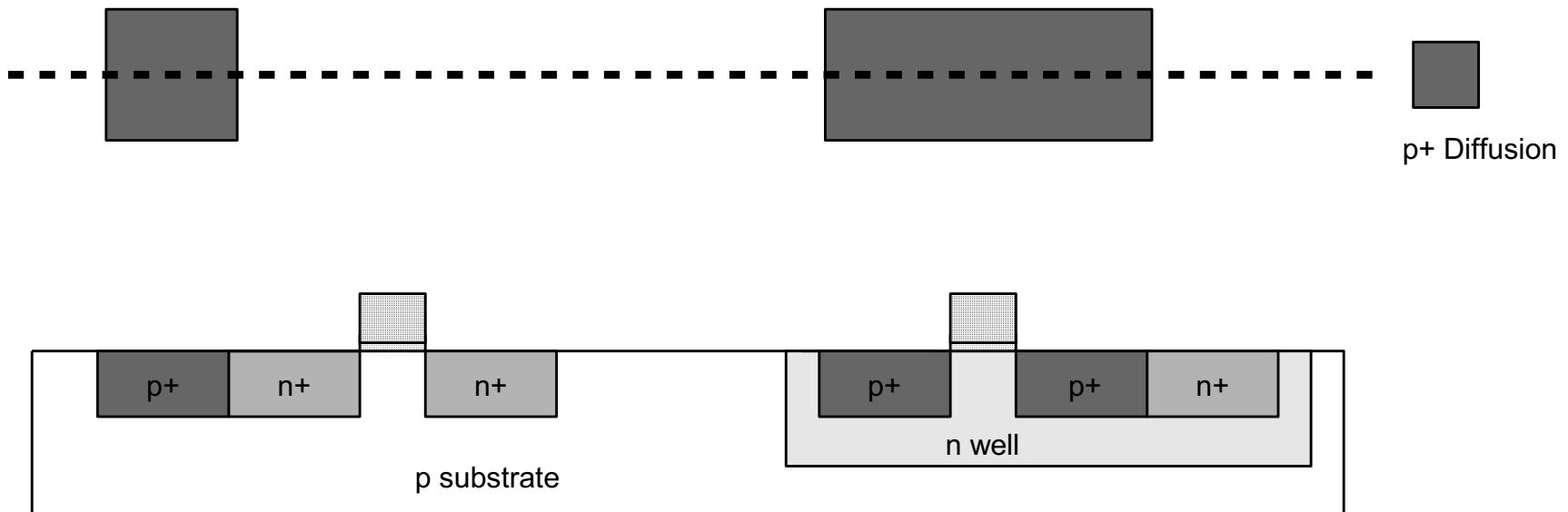
# N-diffusion cont.

- Strip off oxide to complete patterning step



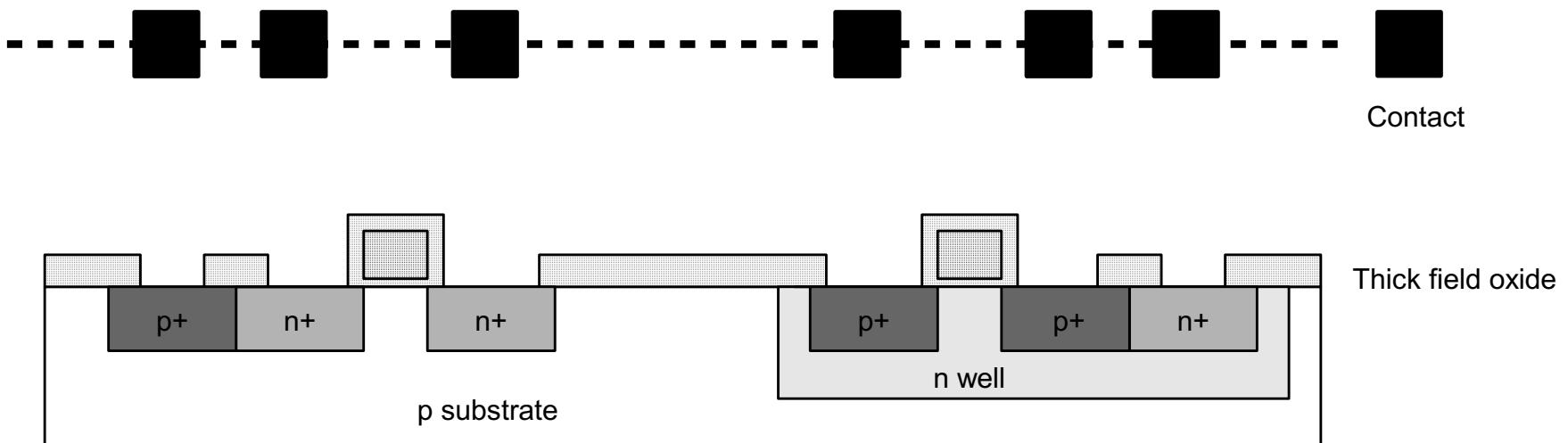
# P-Diffusion

- Similar set of steps form p+ diffusion regions for pMOS source and drain and substrate contact
  - Acceptor atoms (e.g., Boron)



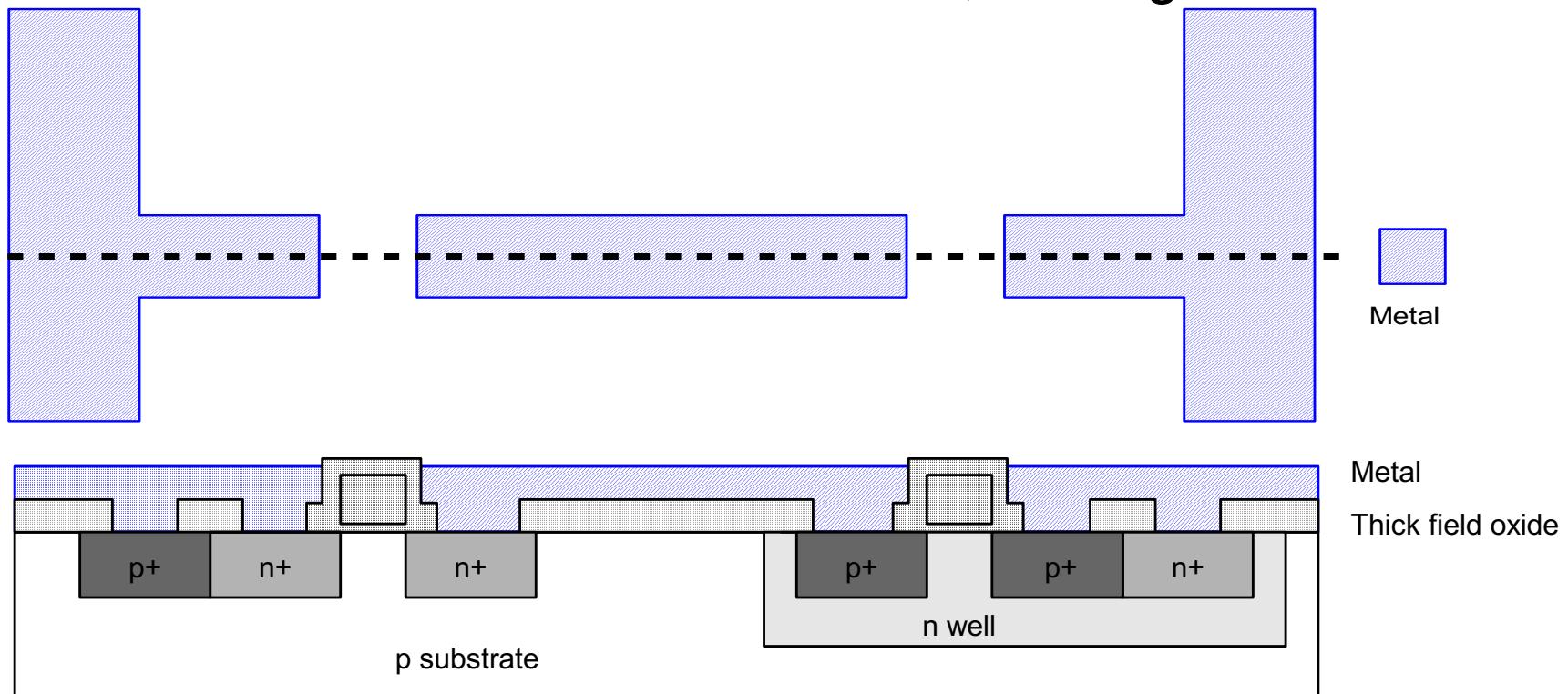
# Contacts

- Now we need to wire together the devices
- Cover chip with thick field oxide (aka field oxide)
- Etch oxide where contact cuts are needed



# Metallization

- ❑ Sputter on aluminum over whole wafer
- ❑ Pattern to remove excess metal, leaving wires



# Feature Size

- Minimum dimensions of masks determine transistor size (and hence speed, cost, and power)
- Feature size  $f$  = distance between source and drain
  - Set by minimum width of polysilicon
- Feature size has improved 30% every 3 years (Moore's law) but no longer is
- Normalize for feature size when describing design rules
- Textbook: expresses rules in terms of  $\lambda = f/2$ 
  - E.g.  $\lambda = 0.3 \mu\text{m}$  in  $0.6 \mu\text{m}$  process

# Design Rules

## In this course's projects:

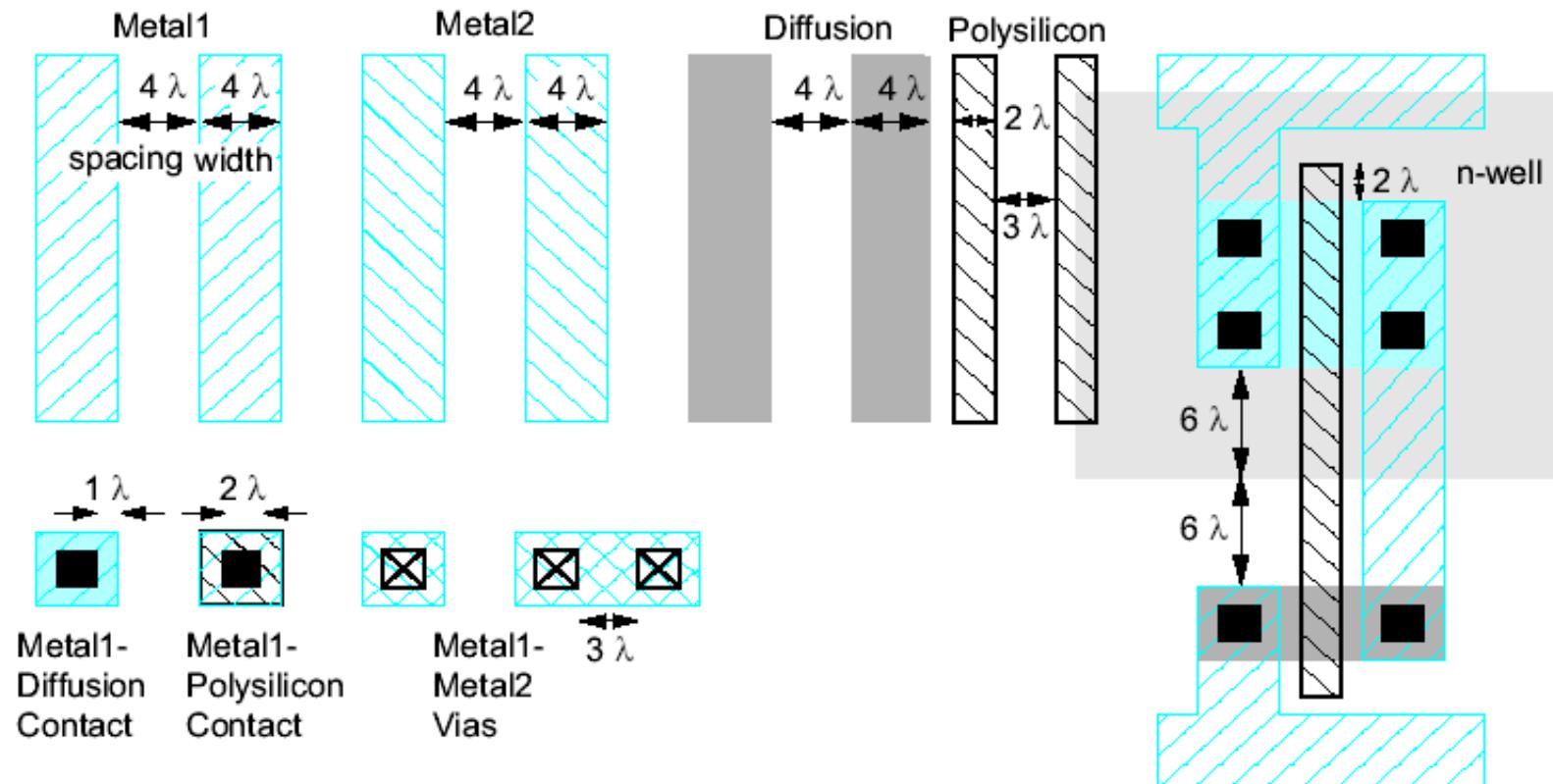
- Individual process design rules/kits in  $\mu\text{m}$ :
  - TSMC 65nm
- Design rules are process-specific and optimized
- Canadian universities: fabrication facilitated by Canadian Microelectronics Corporation (CMC)

## In the textbook:

- Generic design rules valid for groups of technologies based on  $\lambda$
- Design rules can be more conservative as they cover multiple processes each
- USA universities: fabrication facilitated by MOSIS

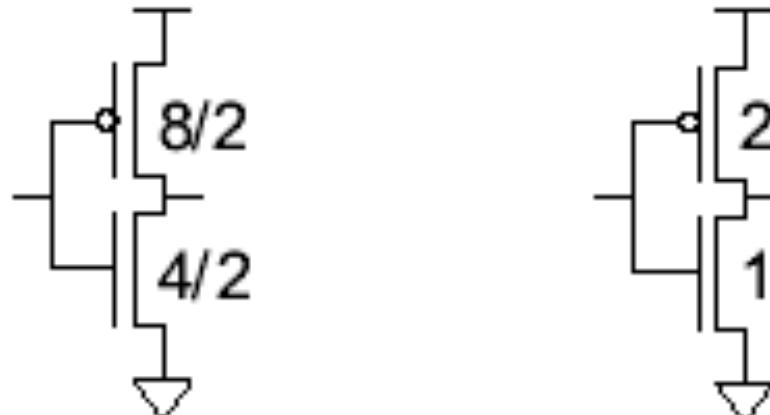
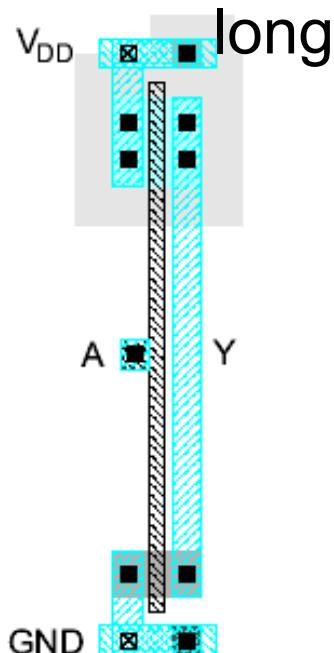
# Simplified Design Rules

- Conservative rules to get you started



# Inverter Layout

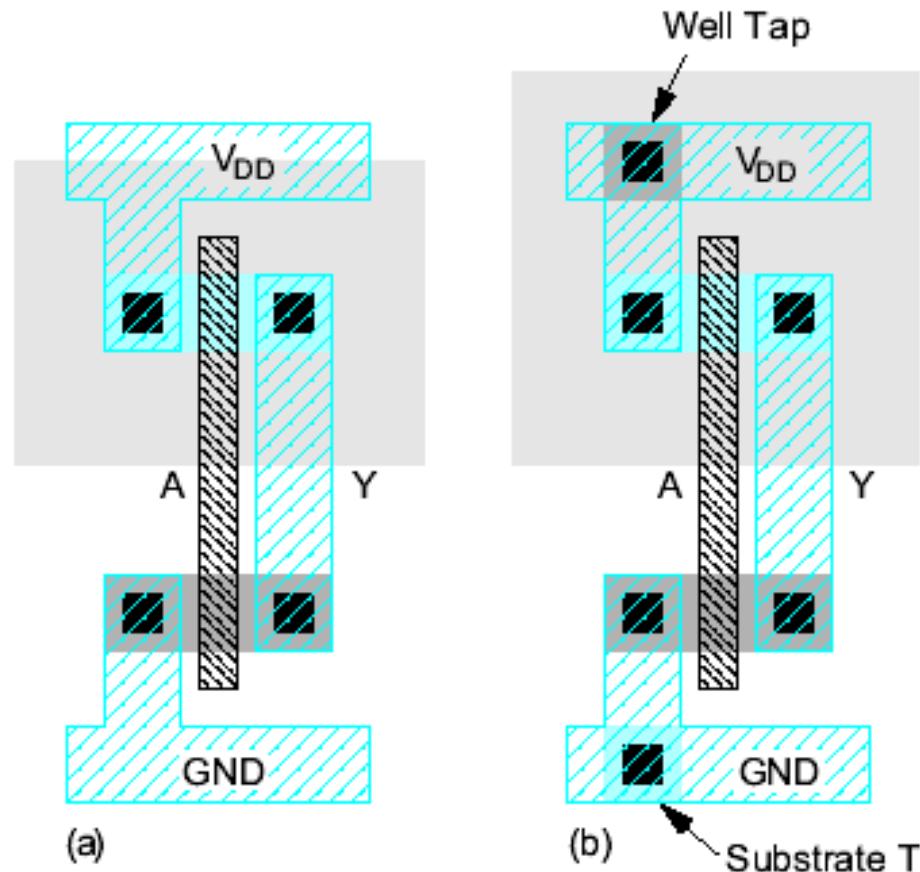
- Transistor dimensions specified as Width / Length
  - Minimum size is  $4\lambda / 2\lambda$ , sometimes called 1 unit
  - In  $f = 0.6 \mu\text{m}$  process, this is  $1.2 \mu\text{m}$  wide,  $0.6 \mu\text{m}$



# Gate Layout

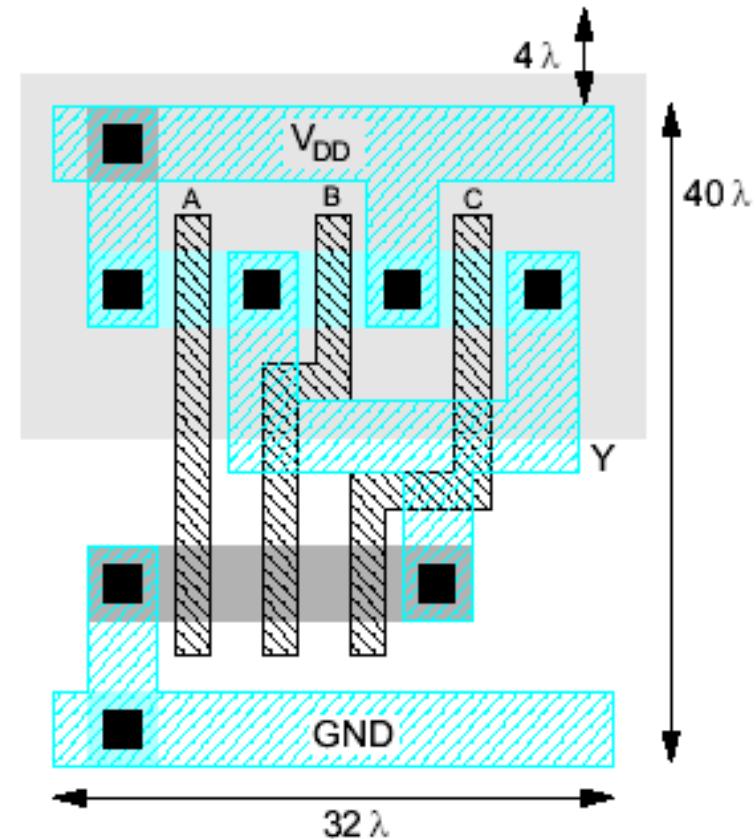
- Layout can be very time consuming
  - Design gates to fit together nicely
  - Build a library of standard cells
  
- Standard cell design methodology
  - $V_{DD}$  and GND should abut (standard height)
  - Adjacent gates should satisfy design rules
  - nMOS at bottom and pMOS at top
  - All gates include well and substrate contacts

# Example: Inverter



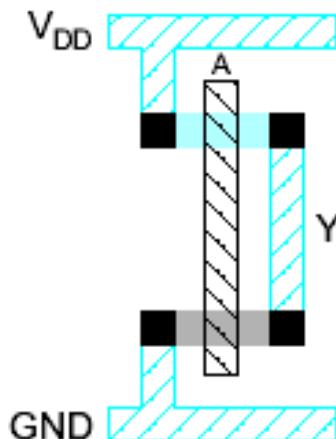
# Example: NAND3

- Horizontal N-diffusion and P-diffusion strips
- Vertical polysilicon gates
- Metal1 V<sub>DD</sub> rail at top
- Metal1 GND rail at bottom
- 32  $\lambda$  by 40  $\lambda$

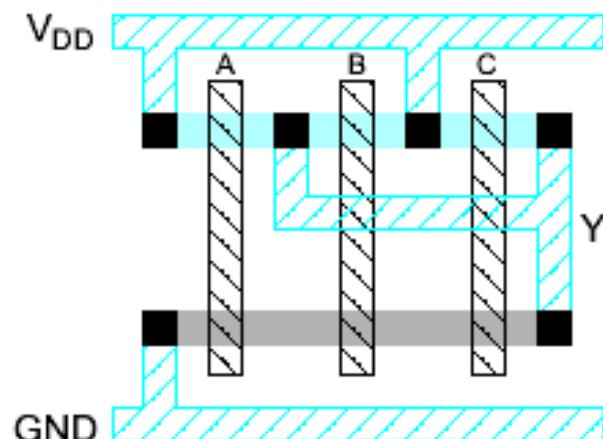


# Stick Diagrams

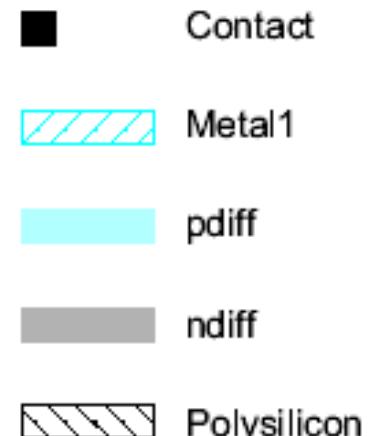
- Stick diagrams help plan layout quickly
  - Need not be to scale
  - Draw with color pencils or dry-erase markers



(a)

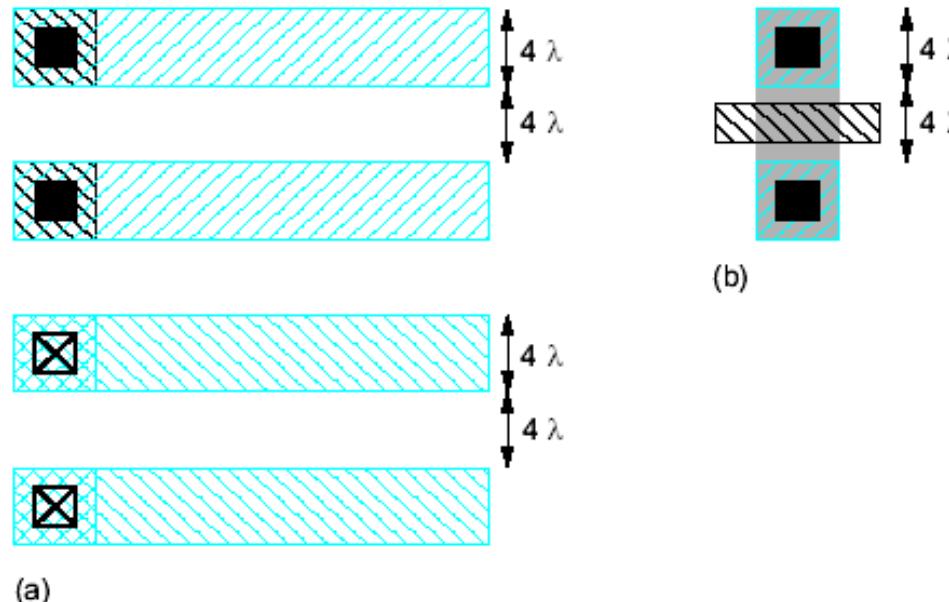


(b)



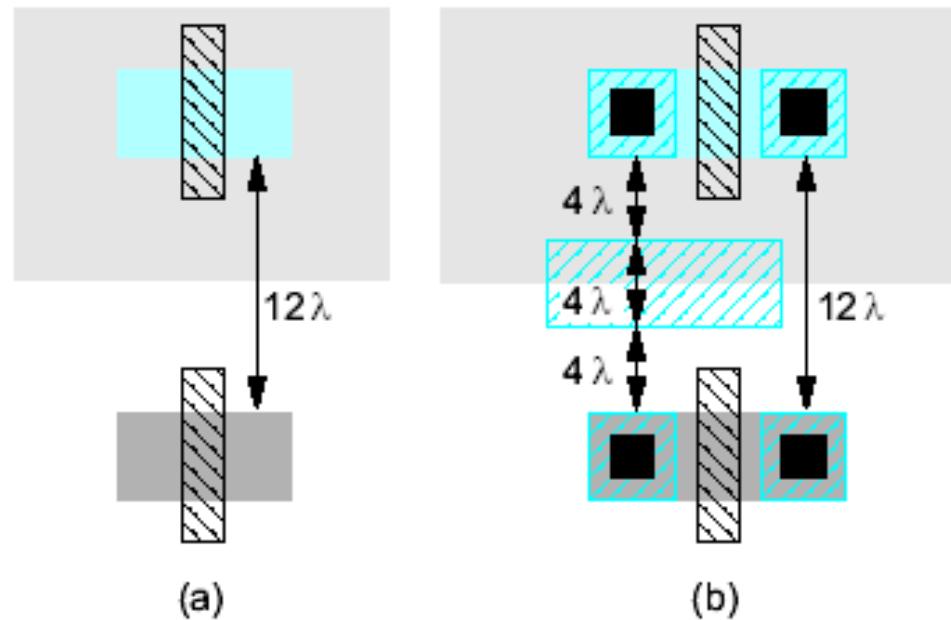
# Wiring Tracks

- A *wiring track* is the space required for a wire
  - $4 \lambda$  width,  $4 \lambda$  spacing from neighbor =  $8 \lambda$  pitch
- Transistors also consume one wiring track



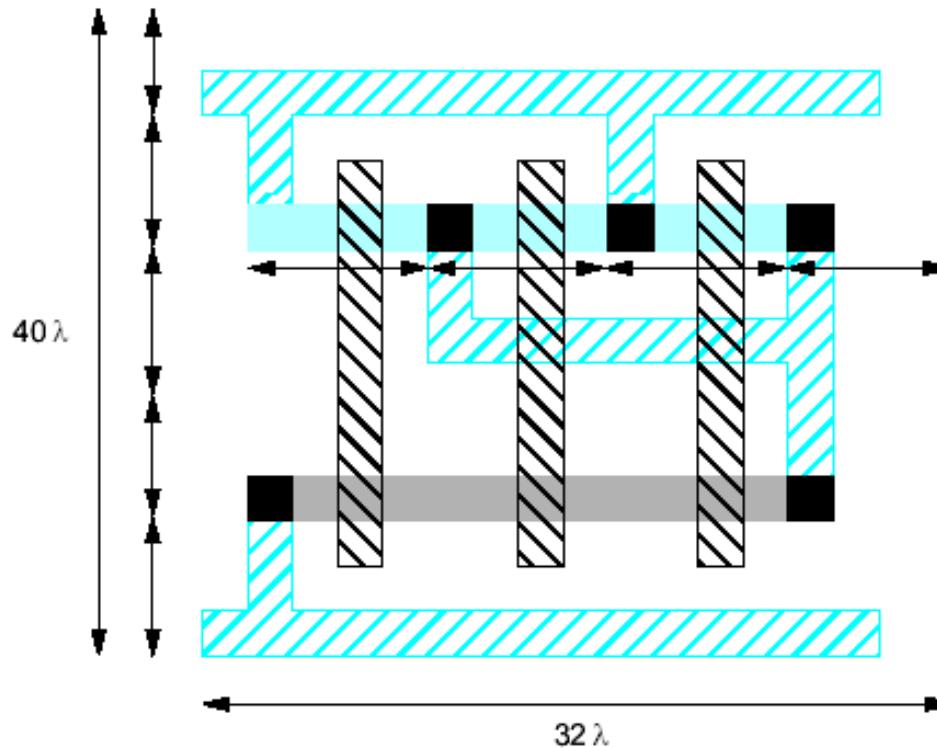
# Well spacing

- Wells must surround transistors by  $6 \lambda$ 
  - Implies  $12 \lambda$  between opposite transistor types
  - Leaves room for one wire track



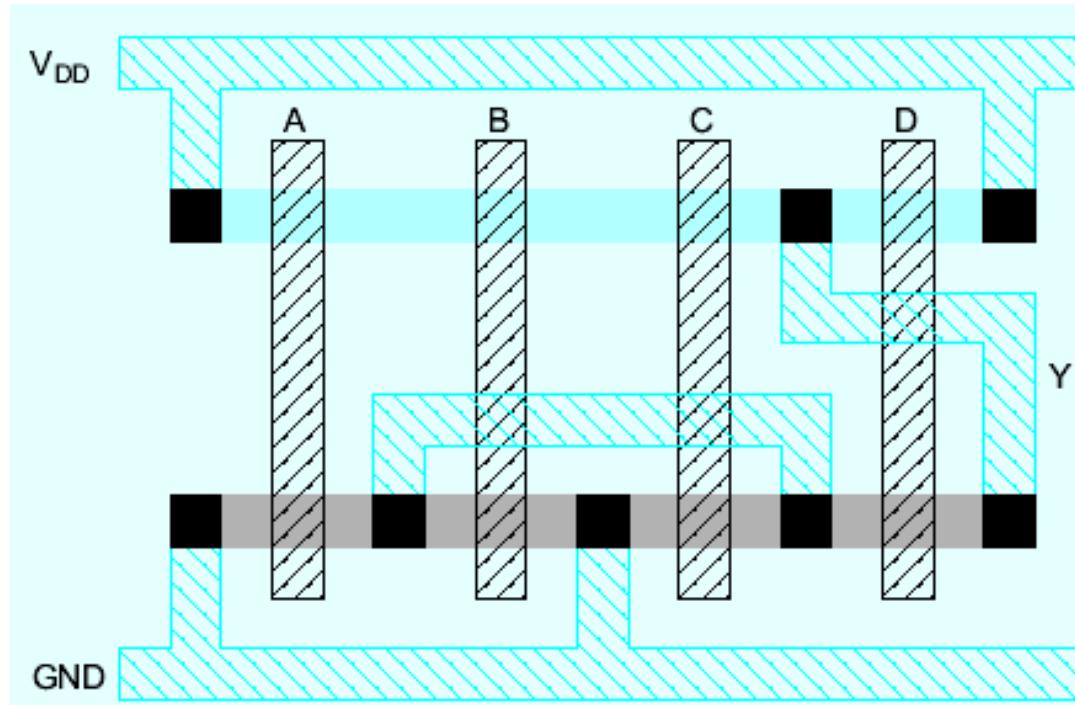
# Area Estimation

- Estimate area by counting wiring tracks
  - Multiply by 8 to express in  $\lambda$



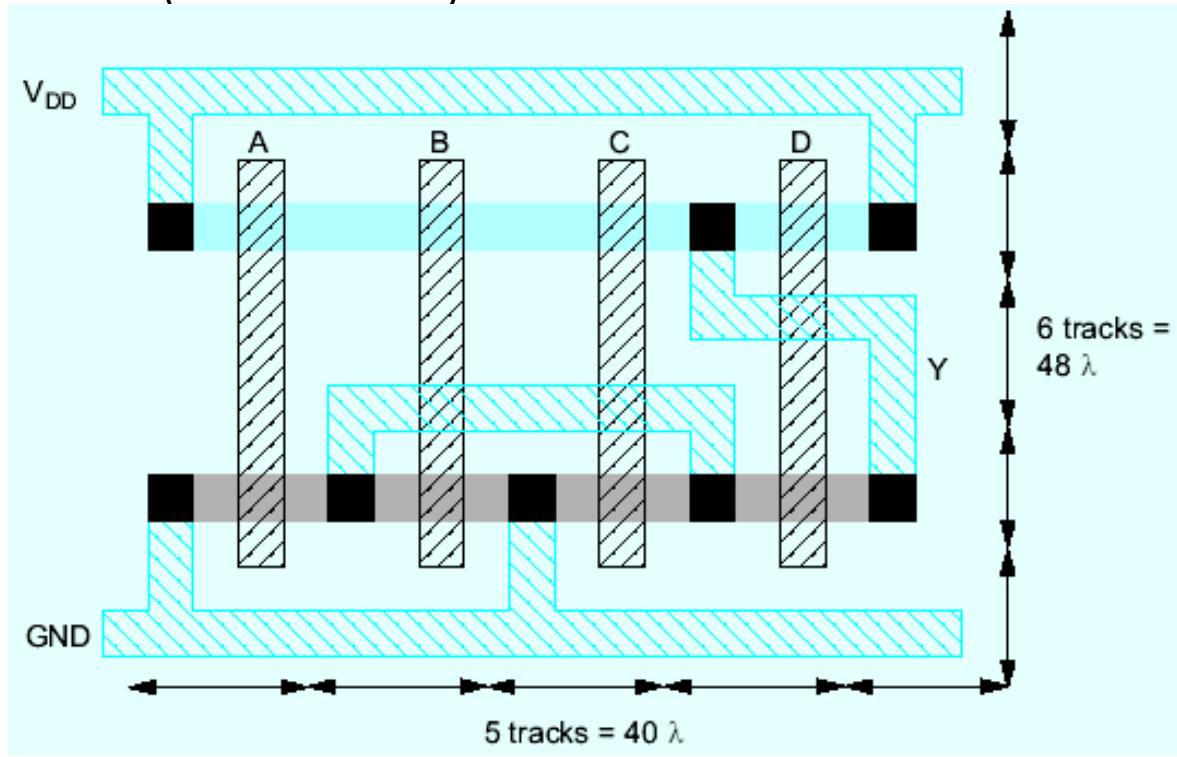
# Example: O3AI

- Sketch a stick diagram for O3AI and estimate area
  - $Y = \overline{(A + B + C)} \cdot D$



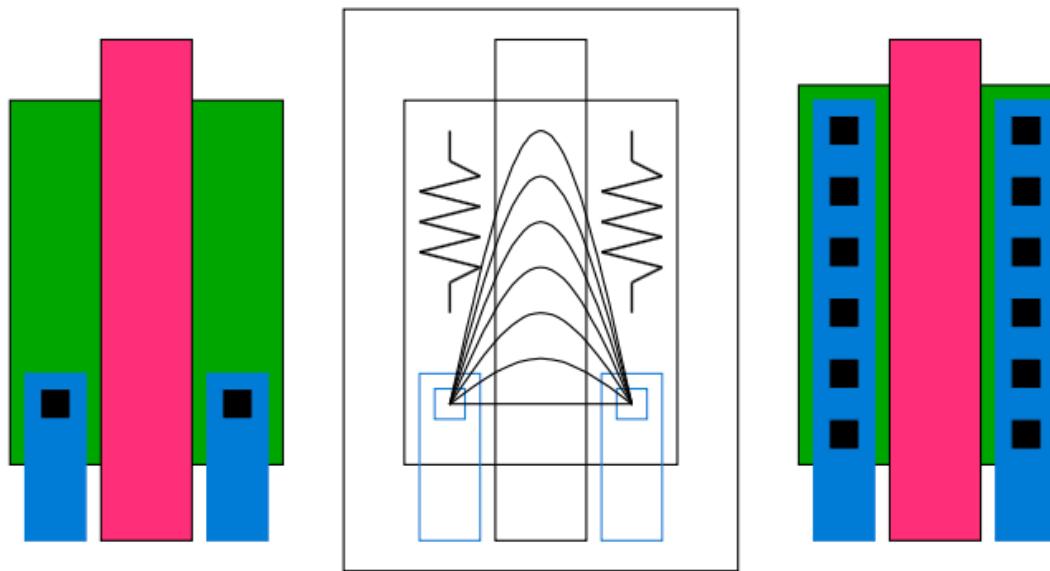
# Example: O3AI

- Sketch a stick diagram for O3AI and estimate area
  - $Y = (A + B + C) \cdot D$



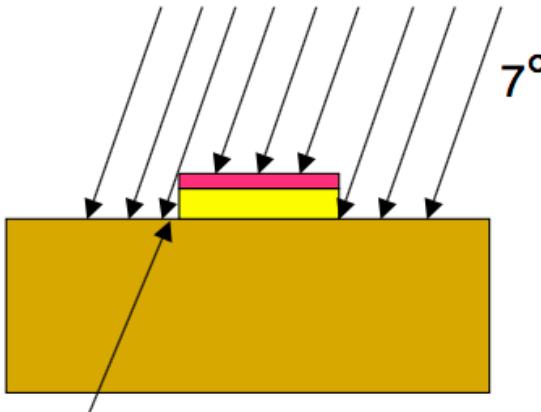
# Source/Drain Connection

- ❖ Ensure good connections

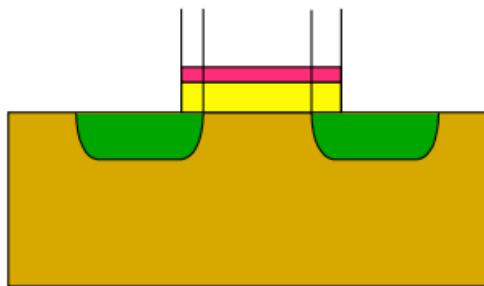


F. Maloberti - Layout of Analog CMOS IC

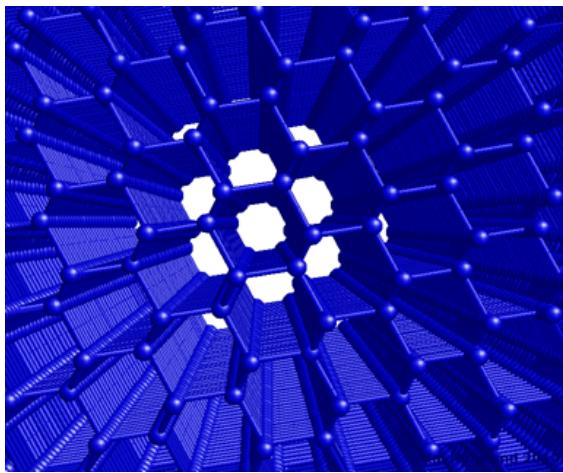
# Asymmetry due to Fabrication



Shadowed region

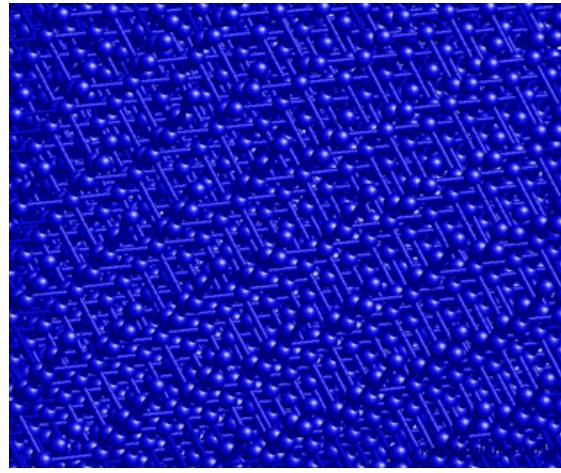


F. Maloberti - Layout of Analog CMOS IC



An about 12 nm thick silicon crystal viewed down the 110 crystal direction.

110: *Miller index*



Same Si crystal viewed from a randomly rotated direction

Wikipedia

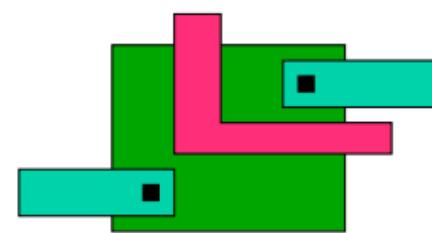
VLSI Design Methodology

- A MOS transistor is not symmetrical:
  - source and drain are not equivalent
- To avoid channeling of implanted ions the wafer is tilted by ~7°
- Channeling is the directionally selective penetration of crystalline solids by a beam of atoms
- Better controlled implant depth and uniformity

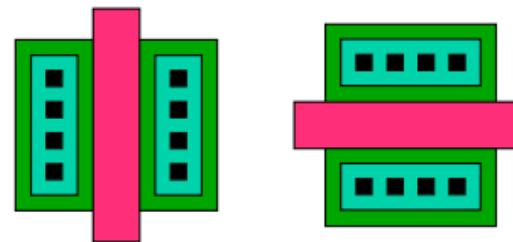
Slide 39

# Matching Single Transistors

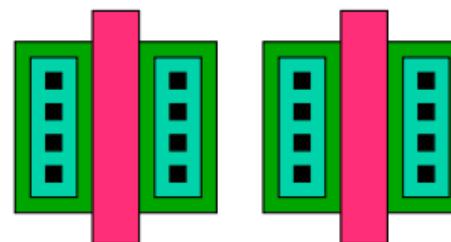
- ❖ Regular (rectangular shape)
  - \* the W and L matter!!



- ❖ Parallel elements
  - \* silicon is unisotropic

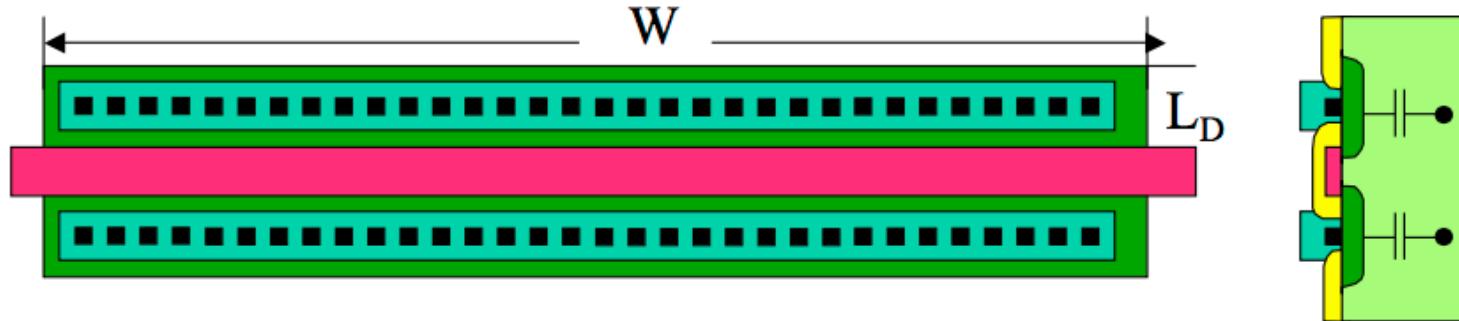


- ❖ Possibly, the current flowing in the same direction



# Parasitics in Transistors

- ❖ Analog transistors often have a large W/L ratio



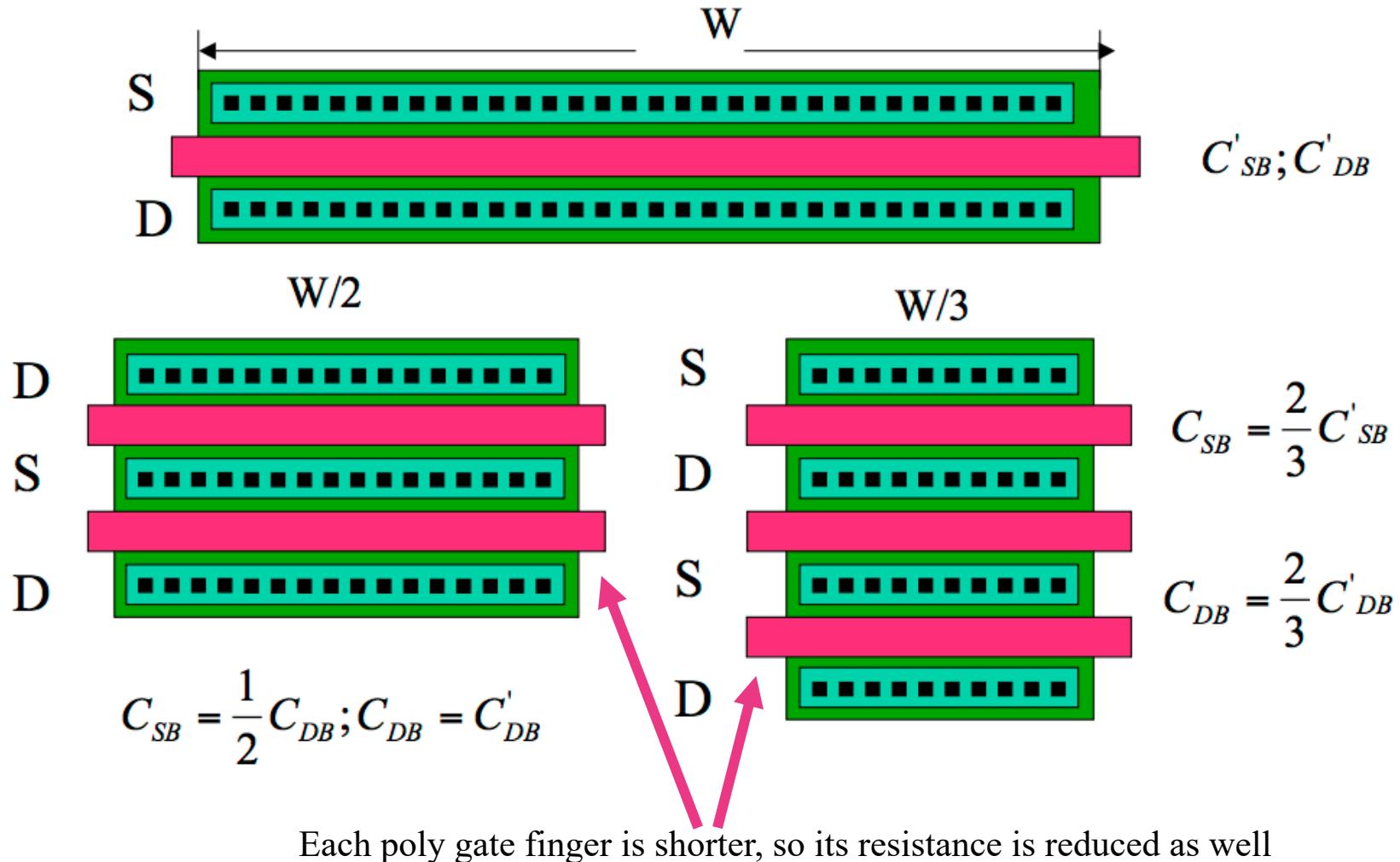
- ❖ Capacitance diffusion substrate

$$C_{SB} = C_{DB} = (W + 2l_{diff})(L_D + 2l_{diff}) \times C_{\text{diff per unit area}}$$

- ❖ Resistance of the poly gate

$$R_{gate} = L_{gate} R_{sq, poly}$$

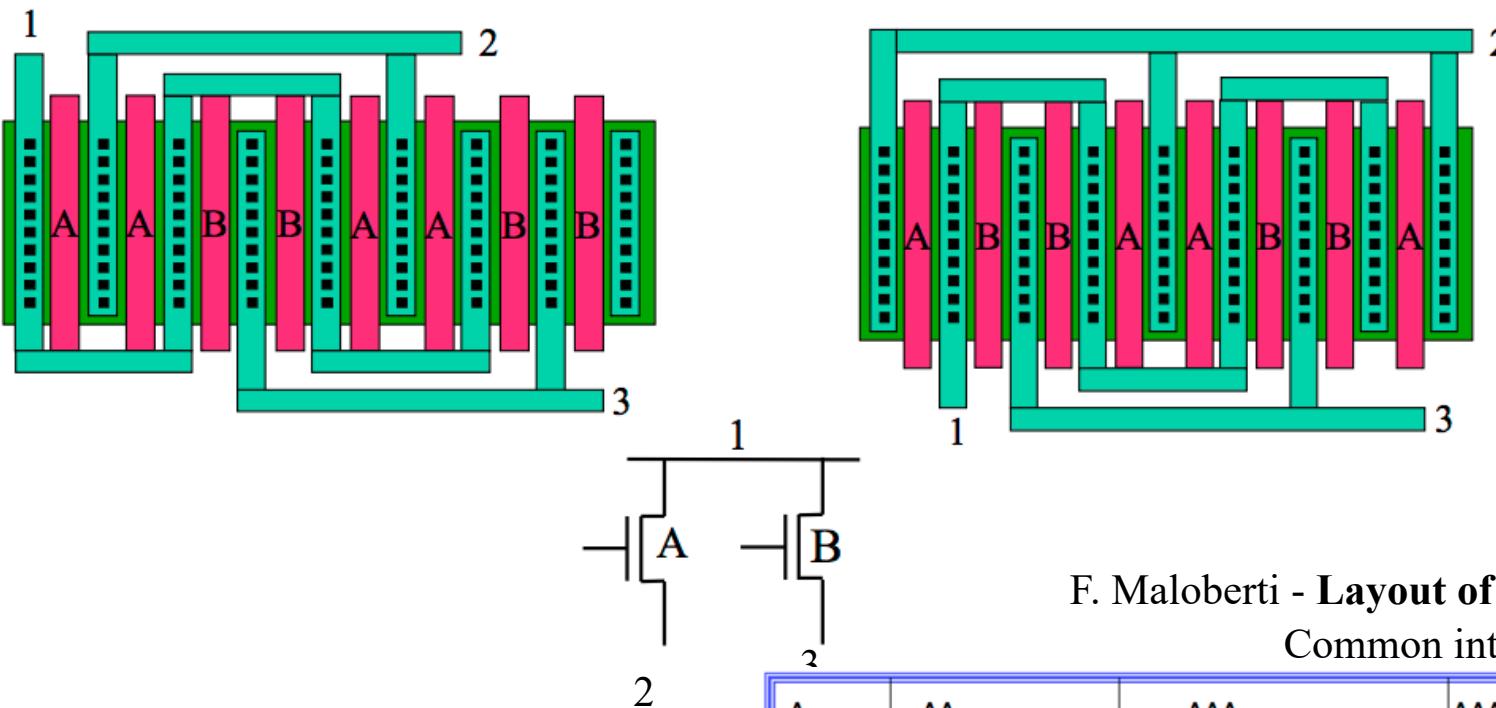
# Use of Multiple Fingers



F. Maloberti - Layout of Analog CMOS IC

# Interdigitated Devices

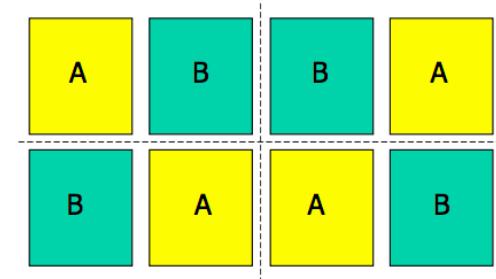
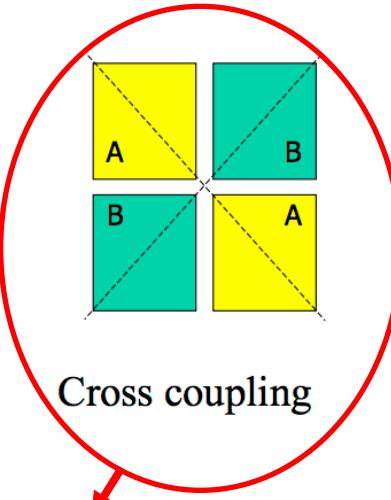
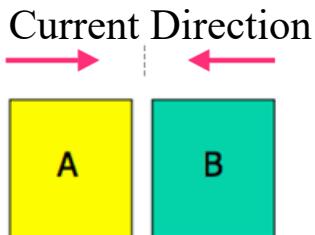
- ❖ Two matched transistors with one node in common
  - ★ split them in an equal part of fingers (for example 4)
  - ★ interdigitate the 8 elements: AABBAABB or ABBAABBA



F. Maloberti - Layout of Analog CMOS IC  
Common interdigitation patterns

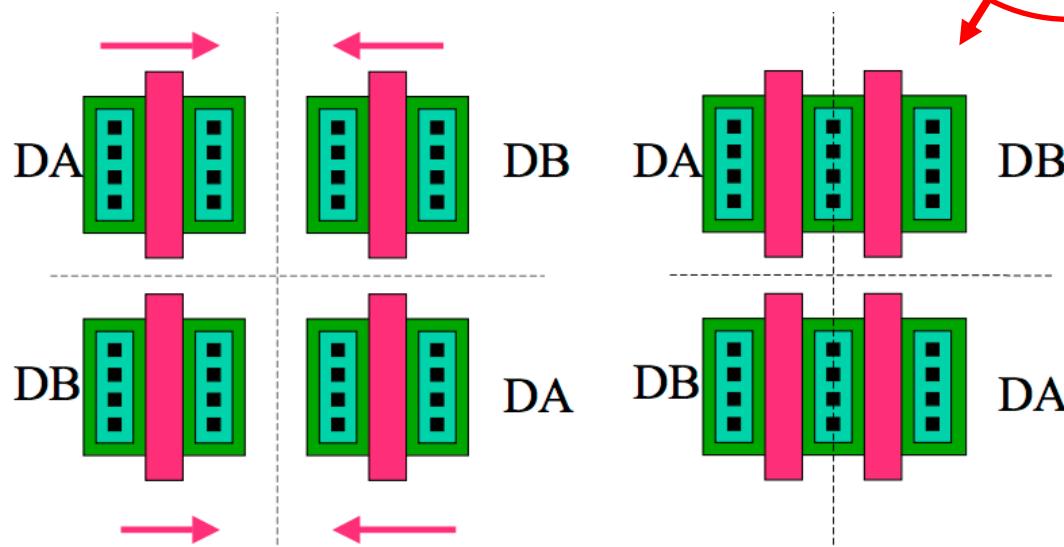
A	AA	AAA	AAAA
AB*	ABBA	ABBAAB*	ABABBABA
ABC*	ABCCBA	ABCACBCA*	ABCABCCBACBA
ABCD*	ABCDDCBA	ABCBCADBCDA*	ABCDDCBAABCDDCBA
ABA	ABAABA	ABAABAABA	ABAABAABAABA
ABABA	ABABAABABA	ABABAABABAABABA	ABABAABABAABABAABABA
AABA*	AABAABAA	AABAAABAAABA*	AABAABAAAABAABAA
AABAA	AABAAAABAA	AABAAAABAAAABAA	AABAAAABAAAABAAAABAA

# Common Centroid Arrays



Simplest case: create a cell with  $\frac{1}{2}$  of the layout and instantiate it twice in mutually-mirrored orientations (e.g., differential pair in opamp, comparator, etc)

Tiling (more sensitive to high-order gradients)



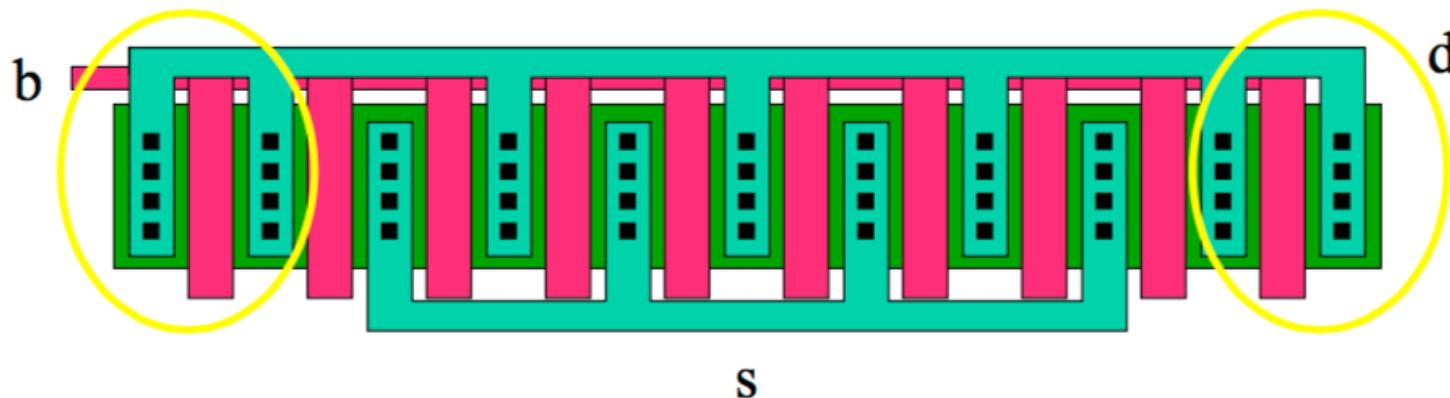
Common centroid patterns			
ABBA BAAB	ABBAABBA BAABBAAB	ABBAABBA BAABBAAB	ABBAABBA BAABBAAB
ABA BAB	ABAABA BABBAB	ABAABA BABBAB	ABAABA BABBAB
ABCCBA CBAABC	ABCCBAABC CBAABCCBA	ABCCBAABC CBAABCCBA	ABCCBAABC CBAABCCBA
AAB BAA	AABBAA BAAAAB	AABBAA BAAAAB	AABBAA BAAAAB

- ❖ Gradients in features are compensated for (at first approximation)
  - \* metal and poly interconnections are more complex

F. Maloberti - Layout of Analog CMOS IC

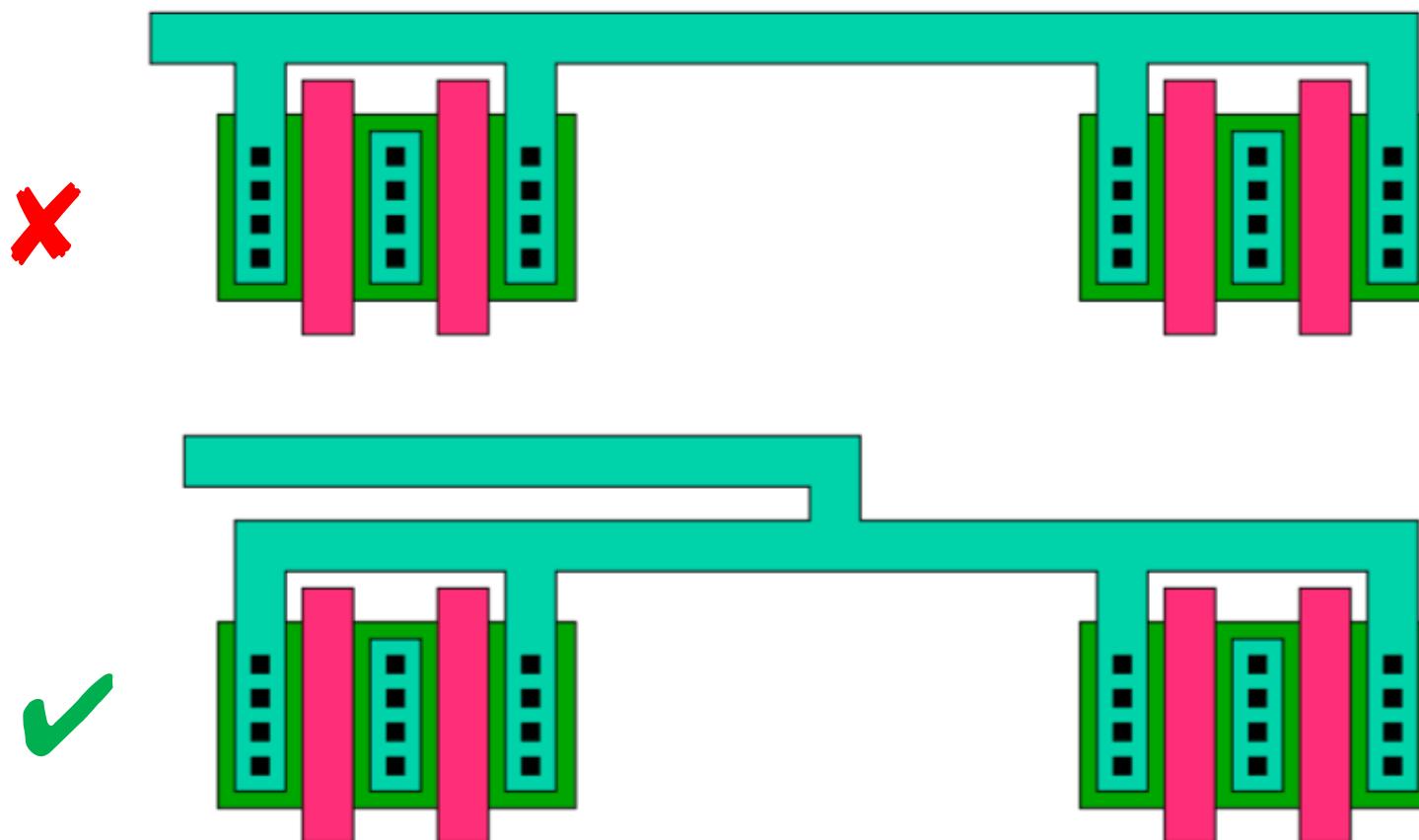
# Dummy Devices on Ends

- ❖ Ending elements have different boundary conditions than the inner elements -> use dummy



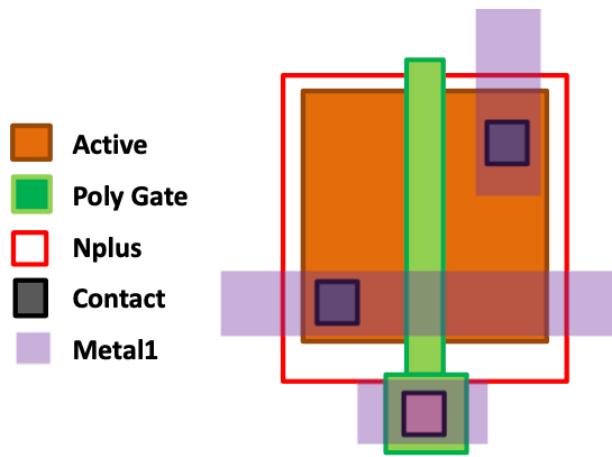
- ❖ Dummies are shorted transistors
  - \* Remember their parasitic contribution!

# Matched Interconnections

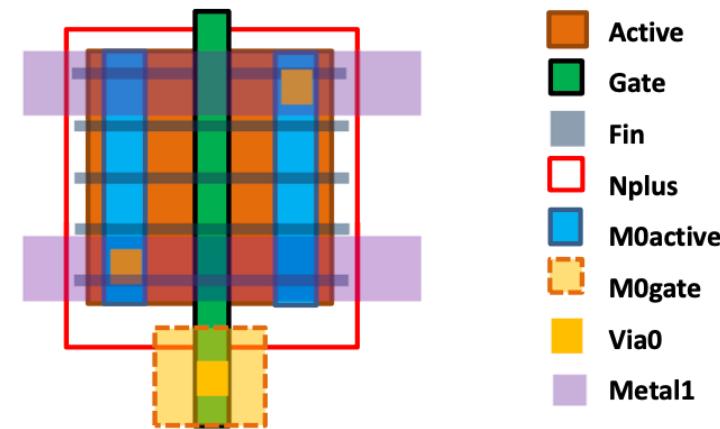


F. Maloberti - Layout of Analog CMOS IC

# Planar Vs FinFET Layout



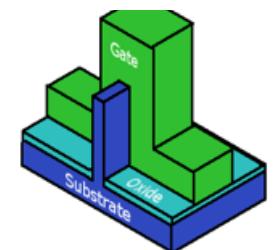
Planar NMOS



FinFET NMOS

- For the planar MOS transistor, we had flexibility in both the transistor channel width and length, as required by the application
- Few, if any, orientation constraints

- A FinFET layout consists of rows of fins in the horizontal direction covered with gates in the vertical direction.
- Orientation and pitch of fins, gates, low-level metals are fixed



“Nanometer CMOS ICs” by Veendrick; and CMC Workshop by TC Carusone