

Advanced VLSI Design

Lecture 4: CMOS Fabrication

Shaahin Hessabi

Department of Computer Engineering

Sharif University of Technology

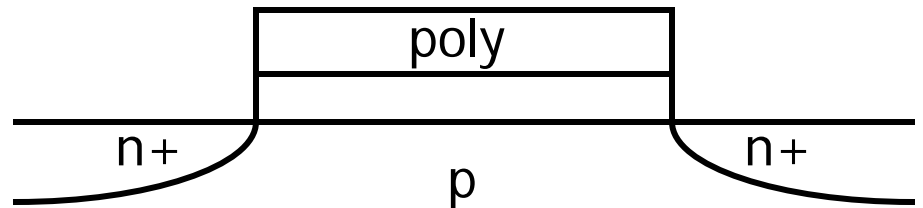
Adapted, with modifications, from lecture notes from
Stanford and Rutgers universities

Overview

- How chips are made
- Design rules for layout

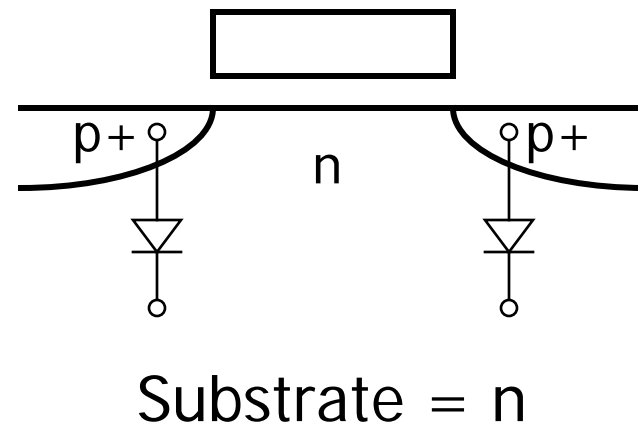
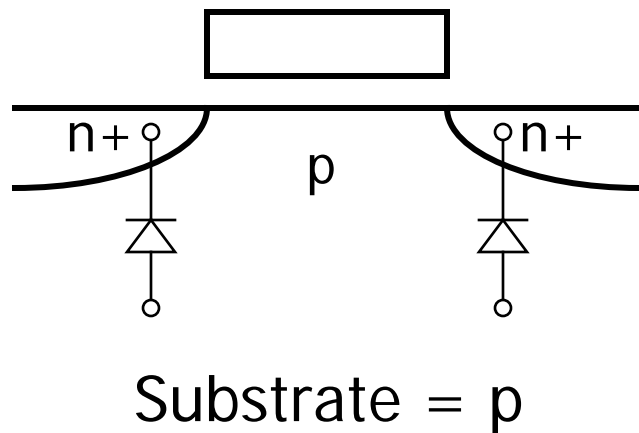
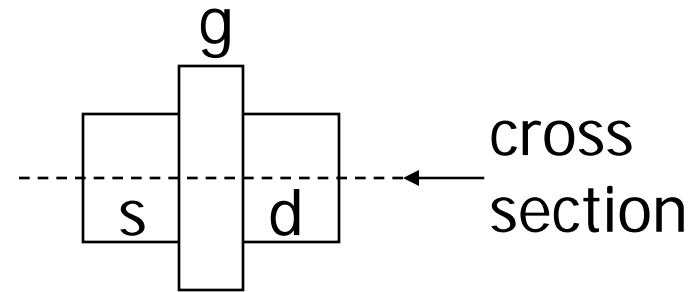
How To Build Transistor

- Diffusion made by adding (diffusing) impurities into silicon
 - n+ (p+) diffusion has lots of impurities (dopants), so higher conductivity
 - p (n) regions lightly doped
 - p region formed first; n+ doped over parts of p region
 - n+ dopant added after poly is done, so that poly blocks dopant



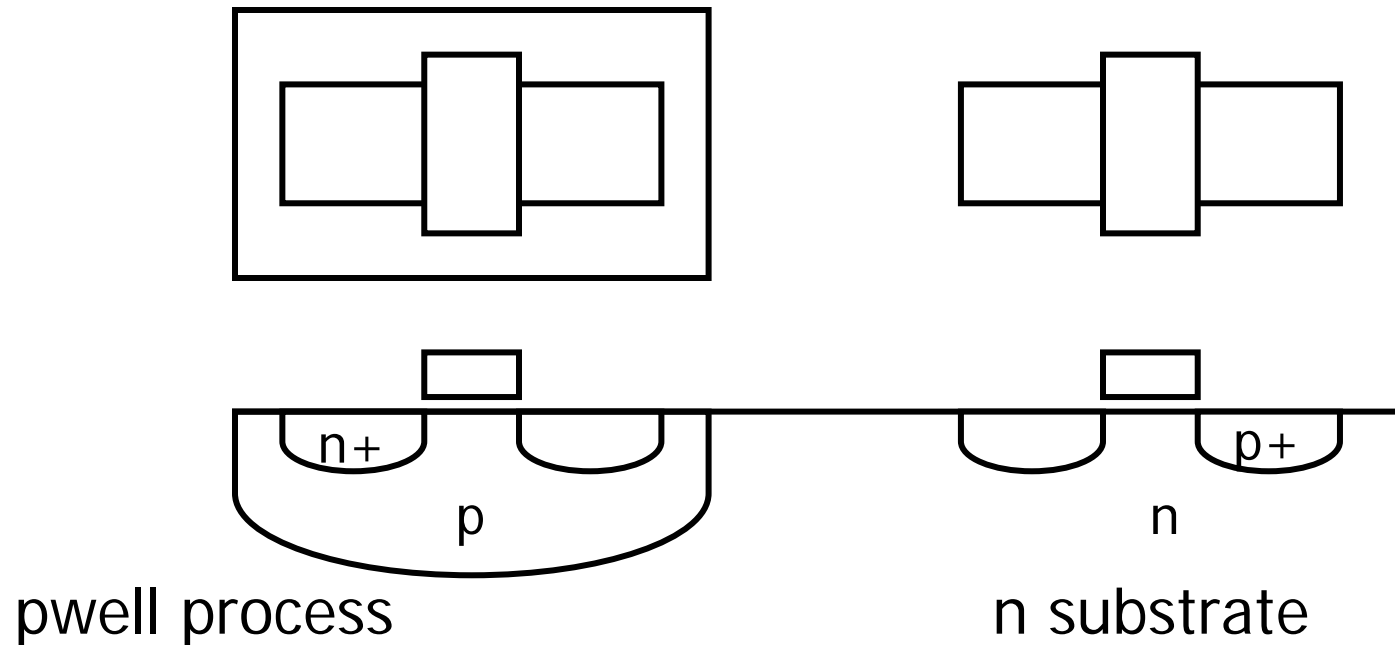
Two Transistor Types

- CMOS requires two types of substrates for isolation of transistors
 - n-type for pMOS
 - p-type for nMOS



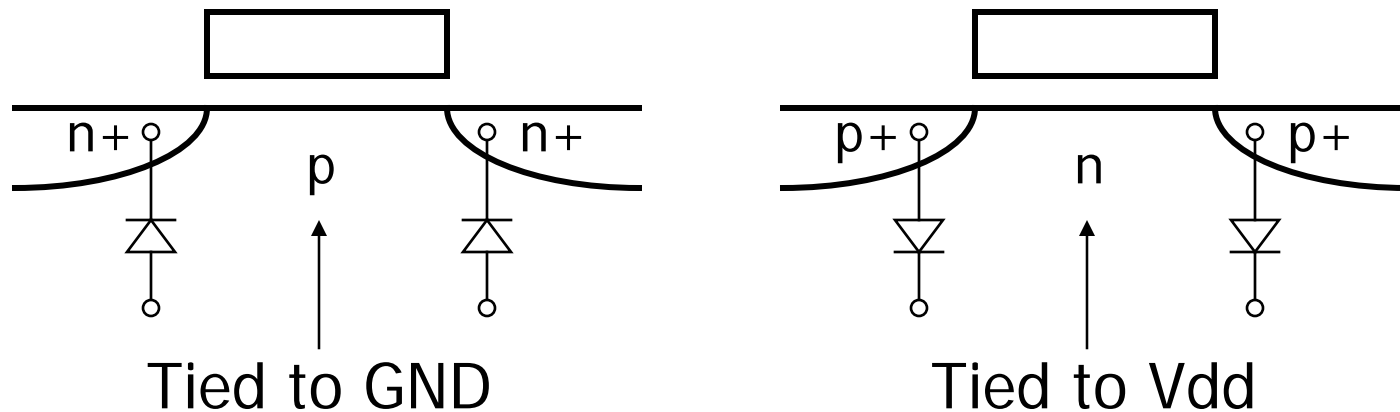
Well: Local Substrate

- Base wafer type may be
 - n-type: add pwell / p-type: add nwell
 - Some have “twin” well



Well Requirement

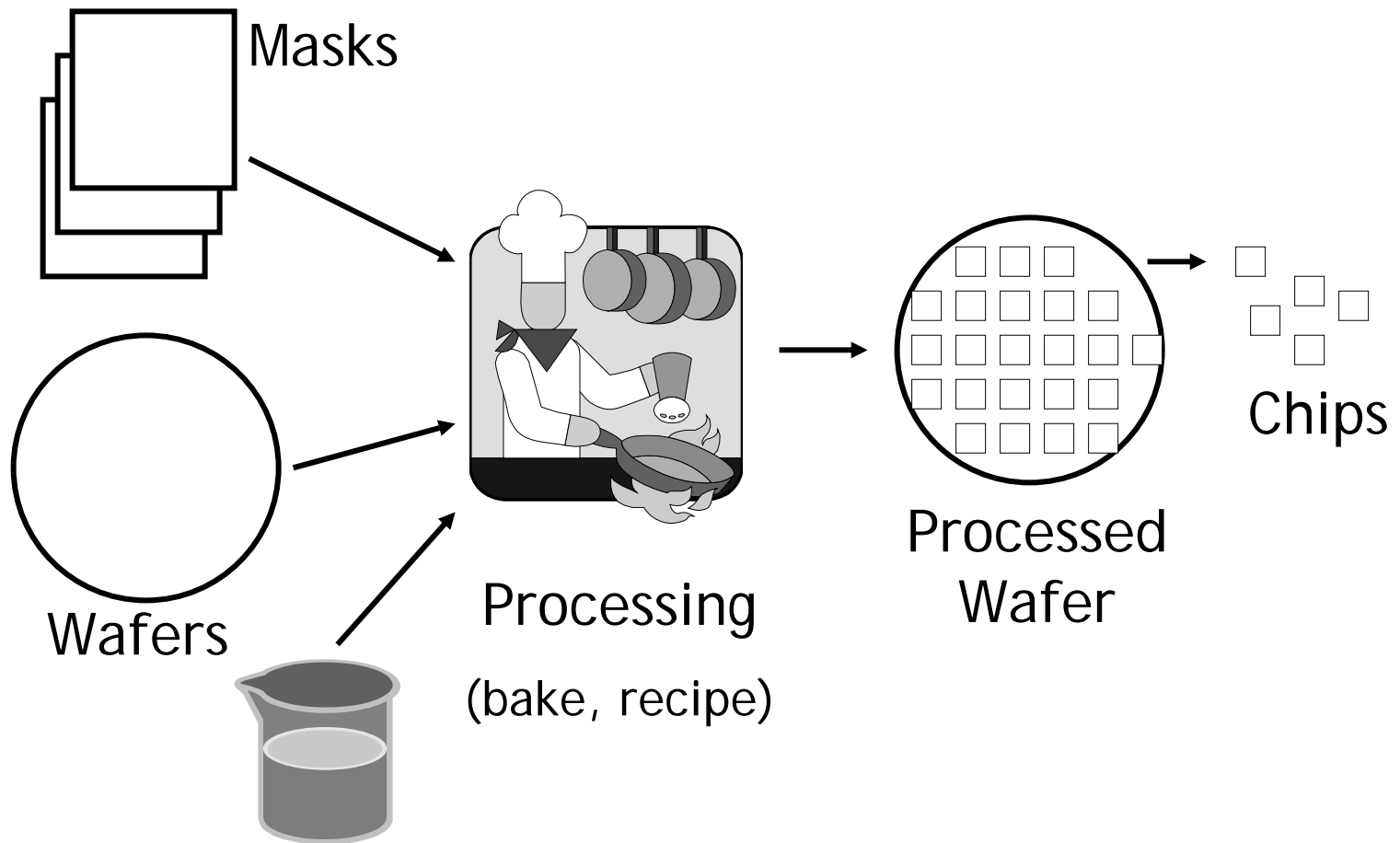
- Well must be tied to a power supply to keep isolation diode reversed biased
 - Using well contacts (ohmic connection to the well)



Well Contacts

- Formed by placing p+ doped region in p-well (n+ region in n-well)
- These regions make good electrical contact to the well (ohmic, not diode), while metal to lightly-doped semiconductor forms poor connection called Schottky Diode
- Well potential equal to the diffusion potential
- Need to have at least one well contact in each well

Fabrication



Basic Fabrication Steps

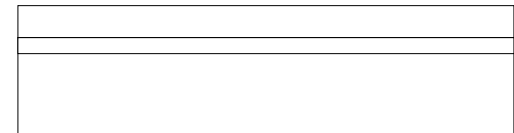
- Transfer image of the design to wafer (photolithography)
- Create layers (diffusion/oxide/metal)
 - Ion implant for diffusion; shoot impurities at silicon
 - Deposition for oxide/metal; usually chemical vapor deposition (CVD)
 - Grow for oxide; place silicon in oxidizing ambient

Basic Processing

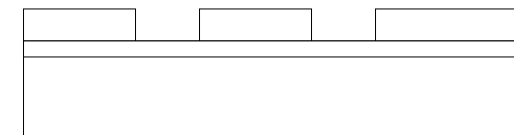
Start with wafer at current step



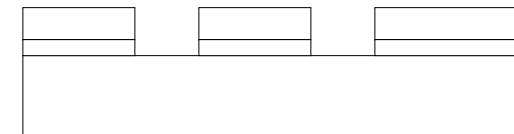
Spin on a photoresist



Pattern photoresist with mask



Step specific processing:
etch, implant, etc...



Wash off resist



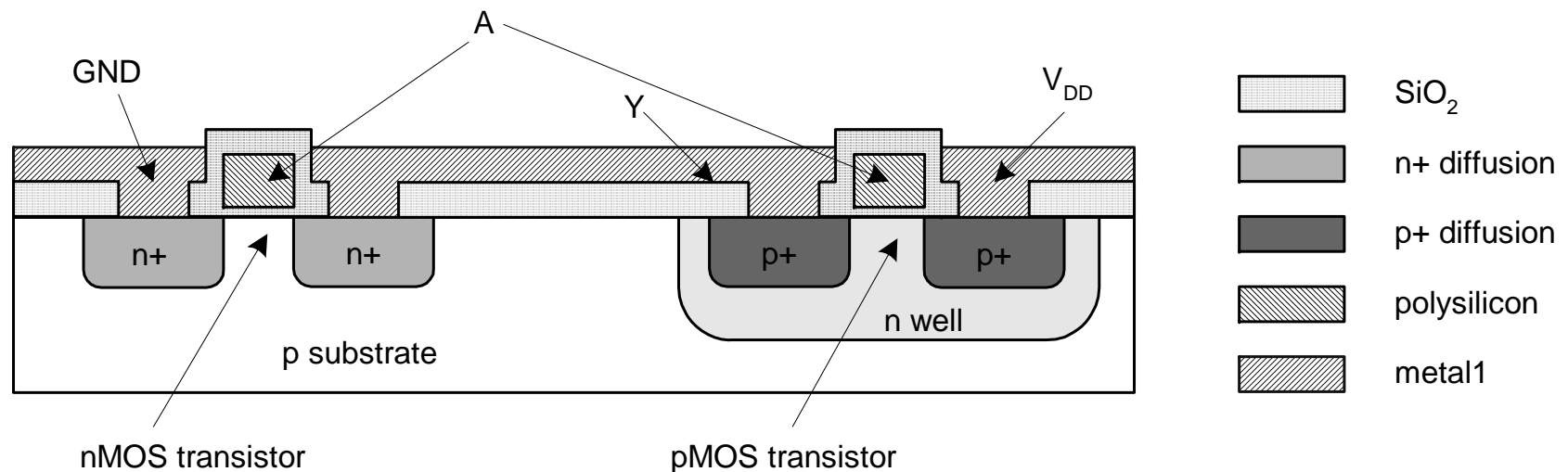
All die on wafer processed in parallel; for some chemical steps, many wafers processed in parallel

Photolithography

- To transfer pattern onto wafer, first need an image to project
 - Glass plate (mask) with image of pattern etched in chrome generated from design database
 - Mask = negative in photography
 - Image optically projected onto wafer using “projection aligner”
 - projection aligner = enlarger in photography
- Mask allows printing on large number of wafers
 - Cost per wafer low, assuming lots of wafers

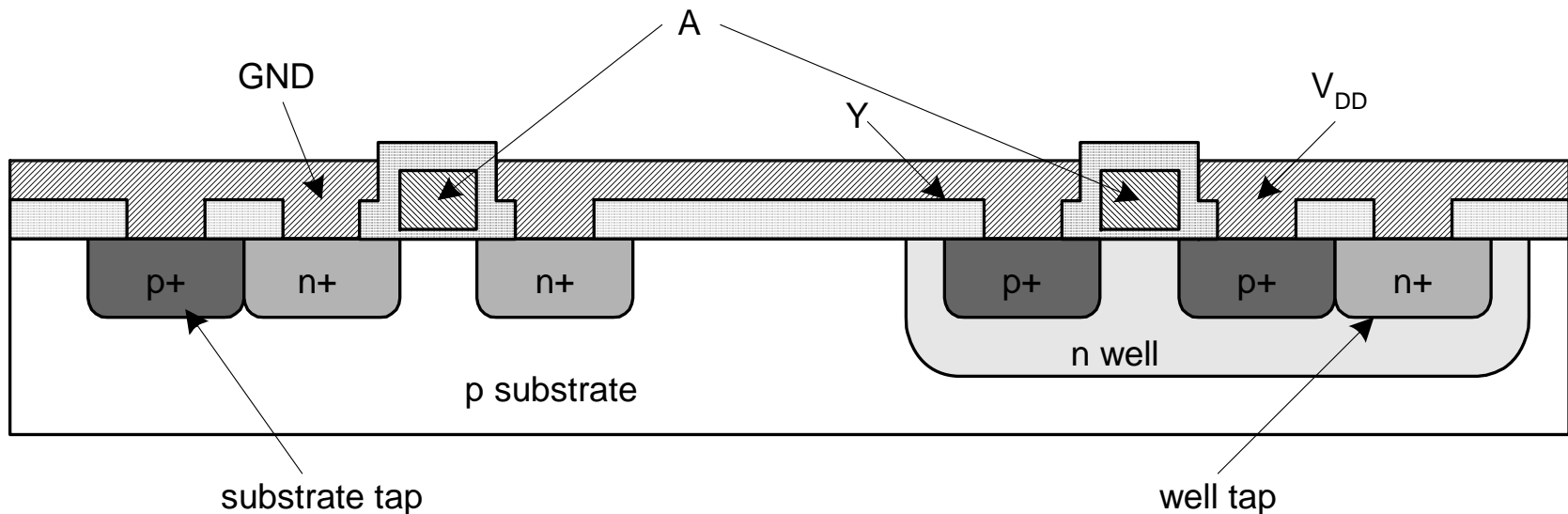
Example: Inverter

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



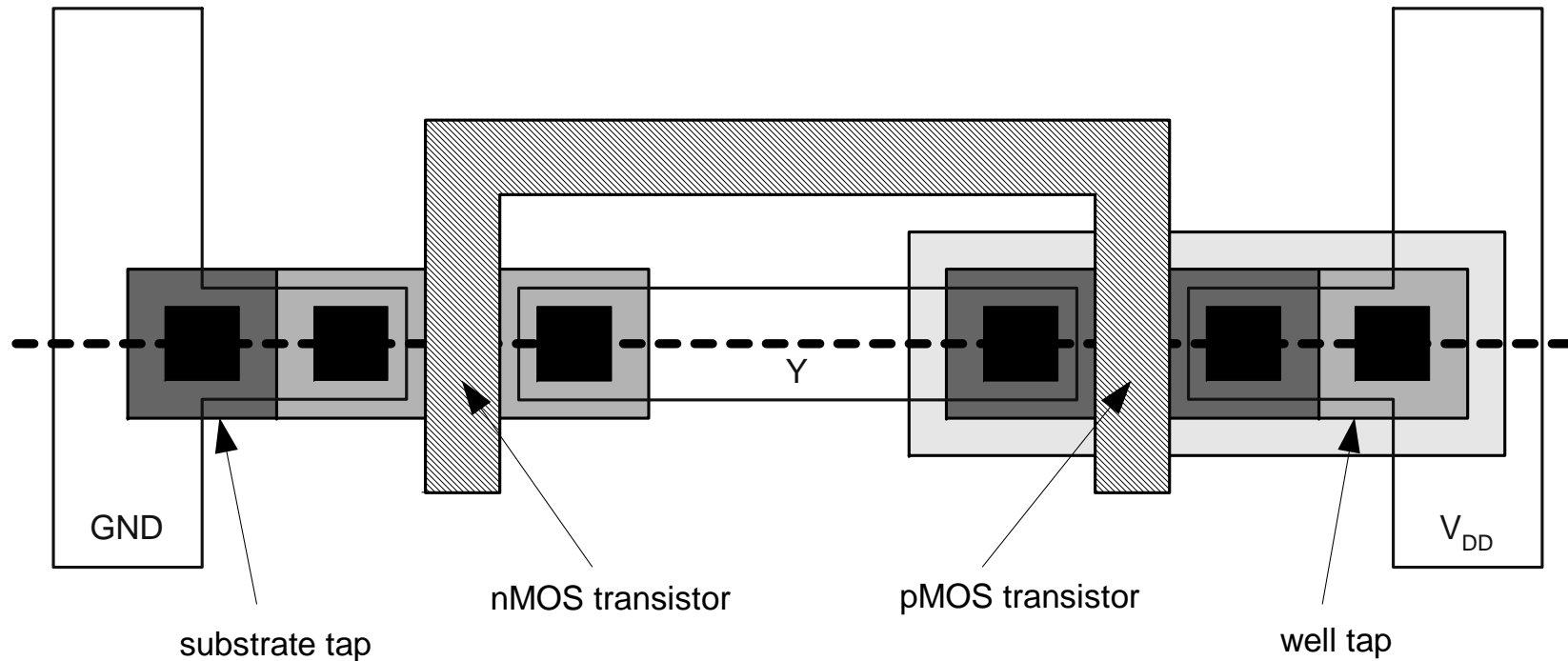
Well and Substrate Taps

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



Inverter Mask Set

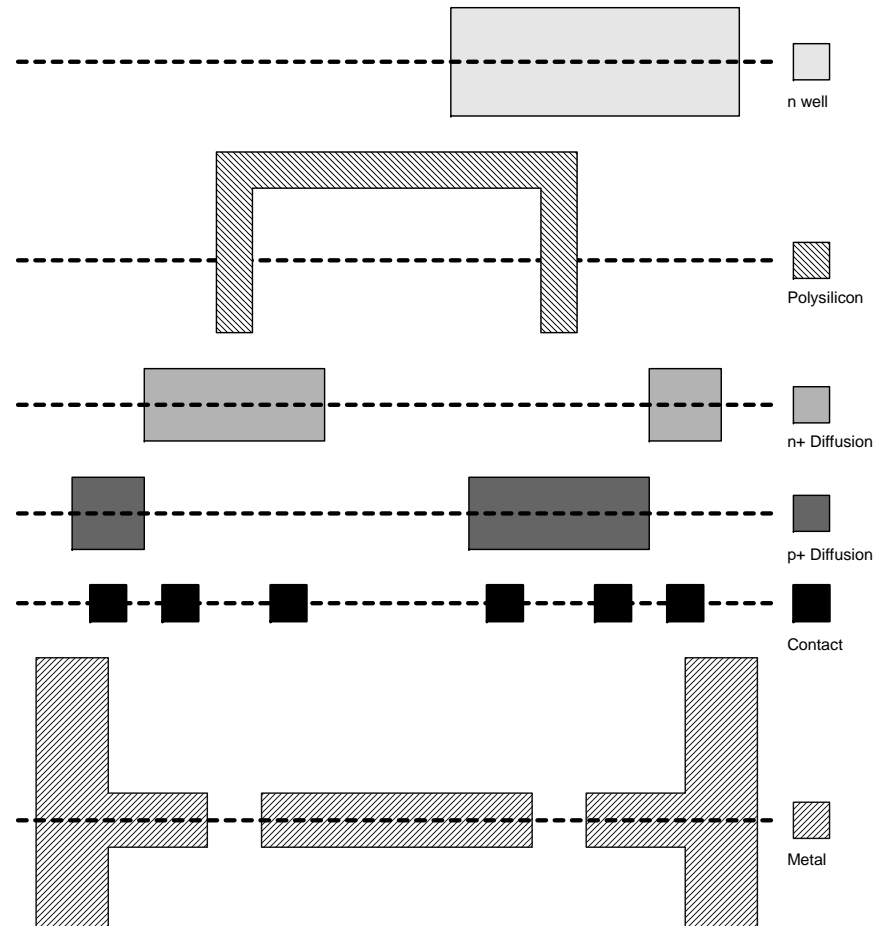
- Transistors and wires are defined by *masks*
- Cross-section taken along dashed line



Detailed Mask Views

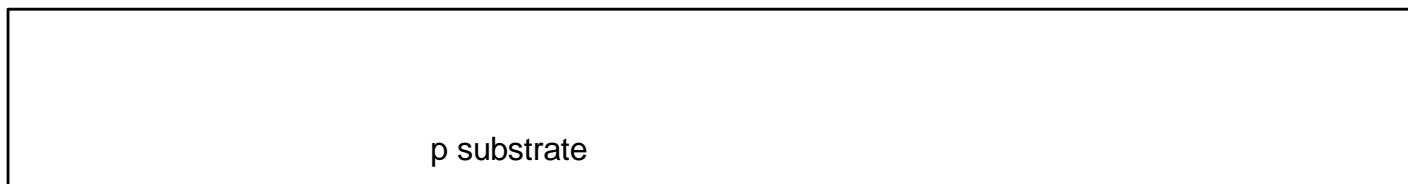
■ Six masks

- 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 SiO_2 (oxide)
 - Remove layer where n-well should be built
 - Implant or diffuse n dopants into exposed wafer
 - Strip off SiO_2



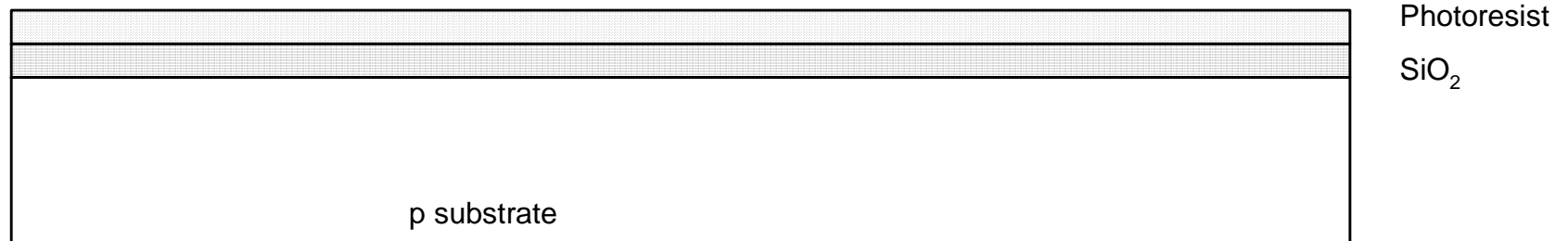
Oxidation

- Grow SiO_2 on top of Si wafer
 - 900 – 1200 C with H_2O or O_2 in oxidation furnace



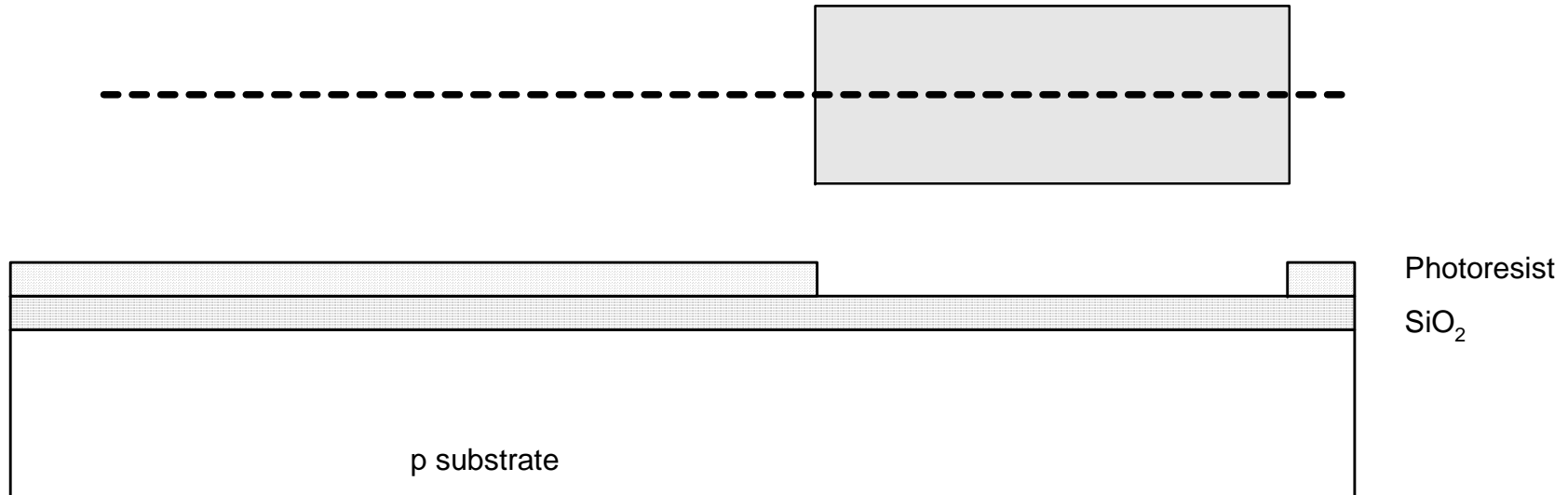
Photoresist

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



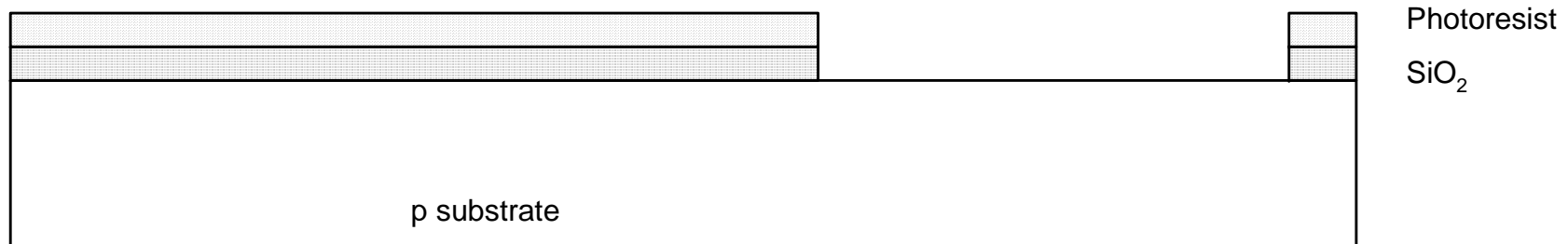
Lithography

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



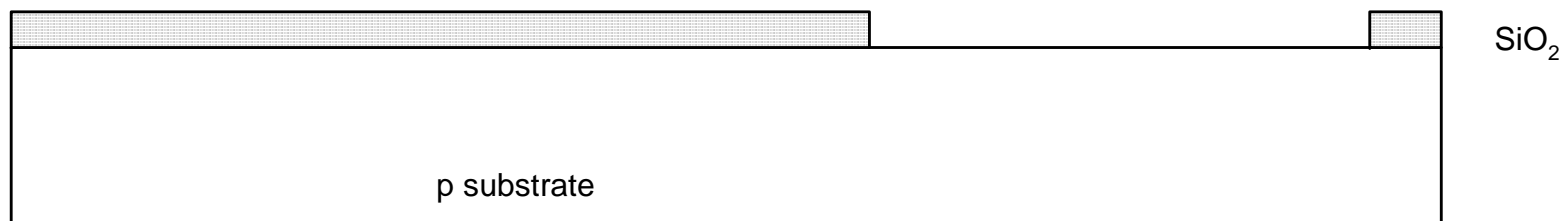
Etch

- Etch oxide with hydrofluoric acid (HF)
- Only attacks oxide where resist has been exposed



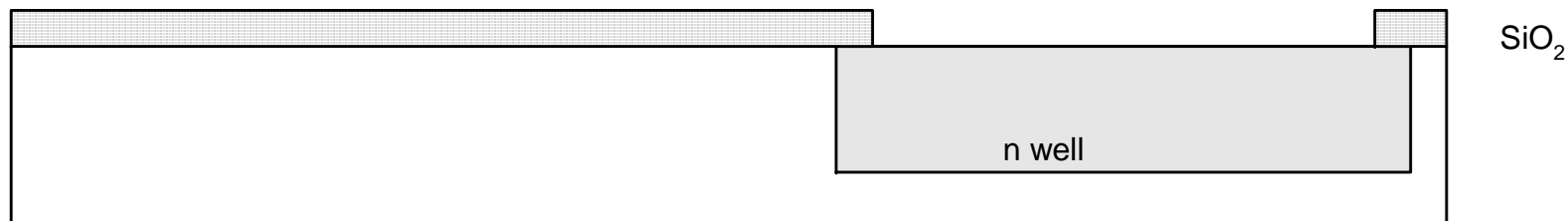
Strip Photoresist

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



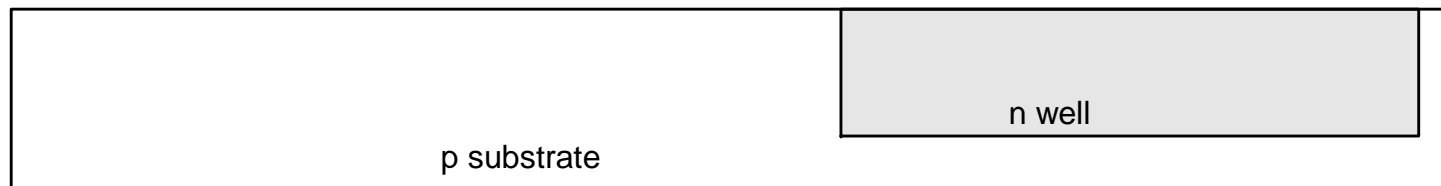
n-Well

- *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 SiO₂, only enter exposed Si



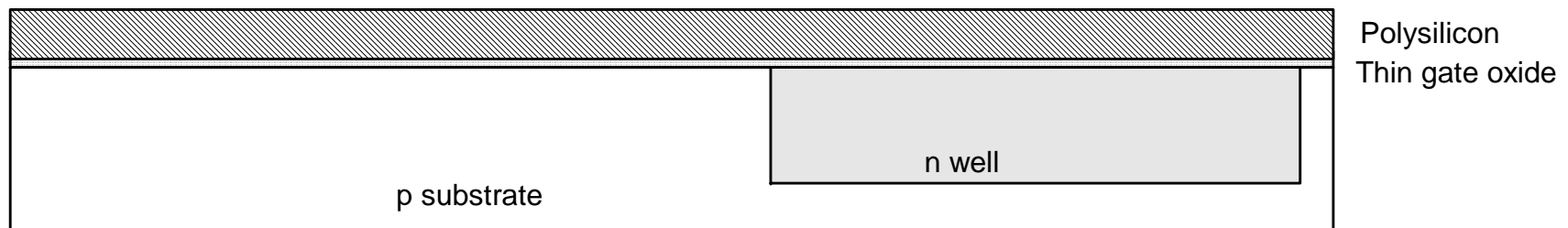
Strip Oxide

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



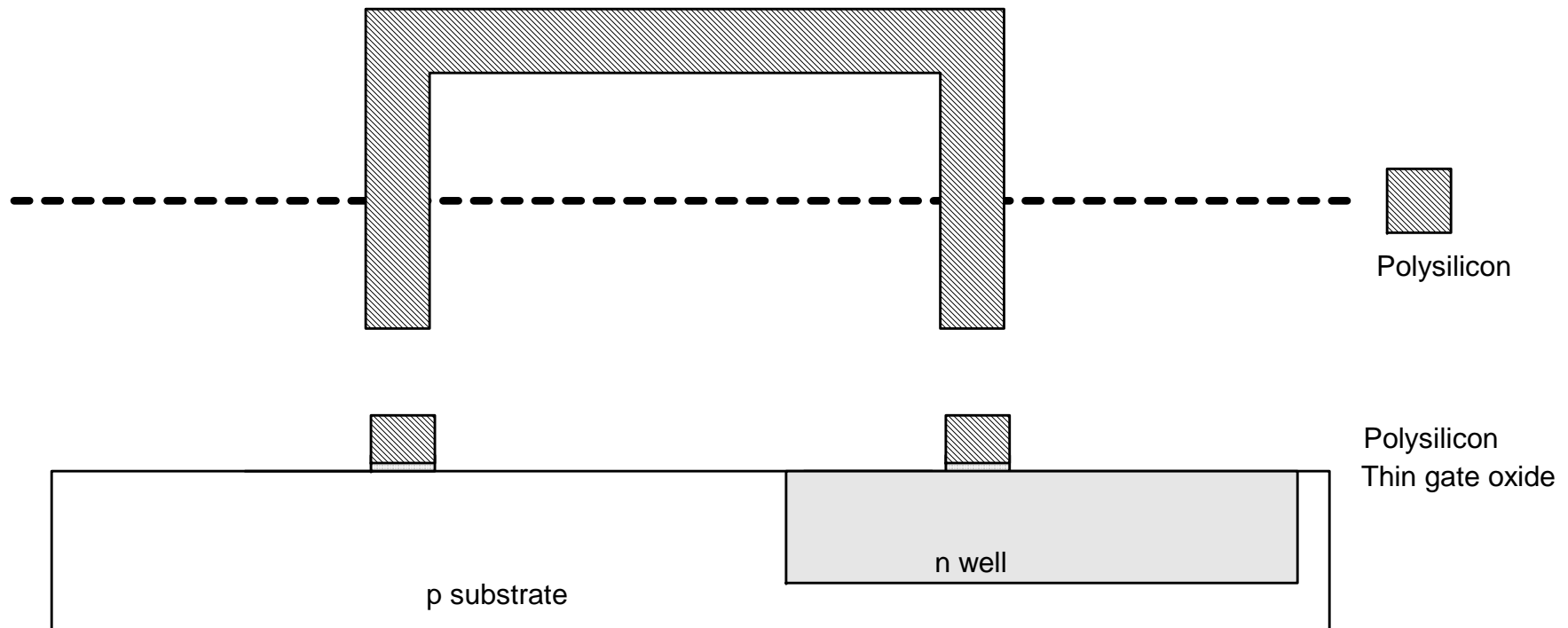
Polysilicon

- Deposit very thin layer of gate oxide
 - $< 20 \text{ \AA}$ (6-7 atomic layers)
- *Chemical Vapor Deposition (CVD)* of silicon layer
 - Place wafer in furnace with Silane gas (SiH_4)
 - Forms many small crystals called polysilicon
 - Heavily doped to be good conductor



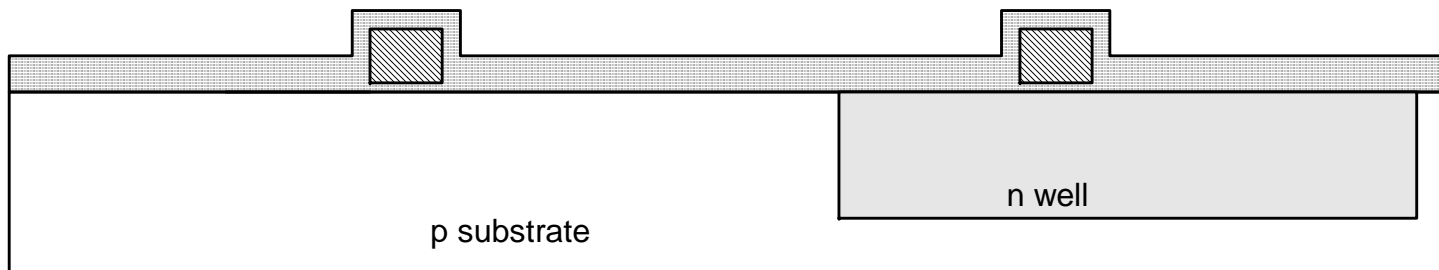
Polysilicon Patterning

- Use same lithography process to pattern polysilicon



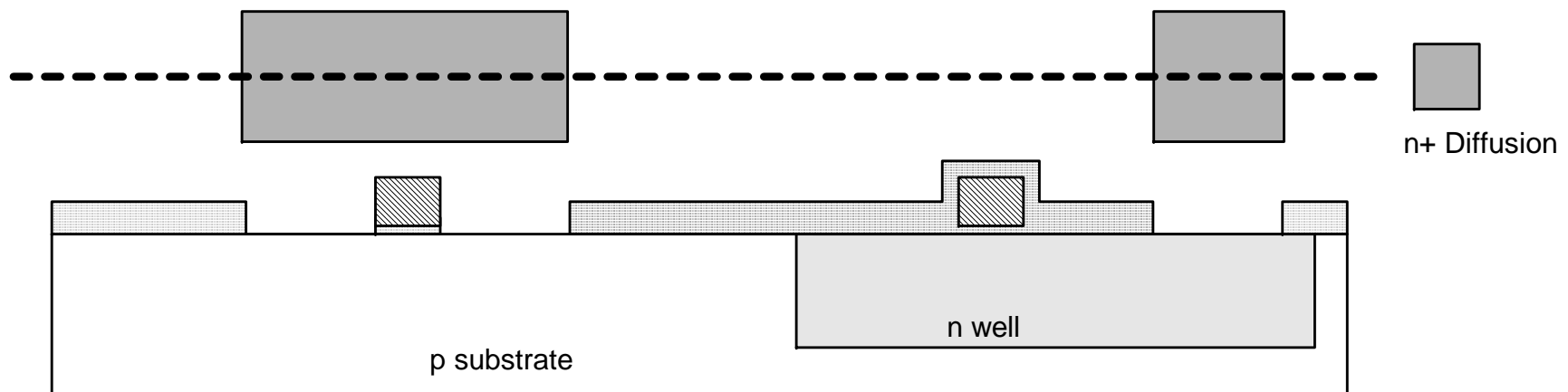
Self-Aligned Process

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



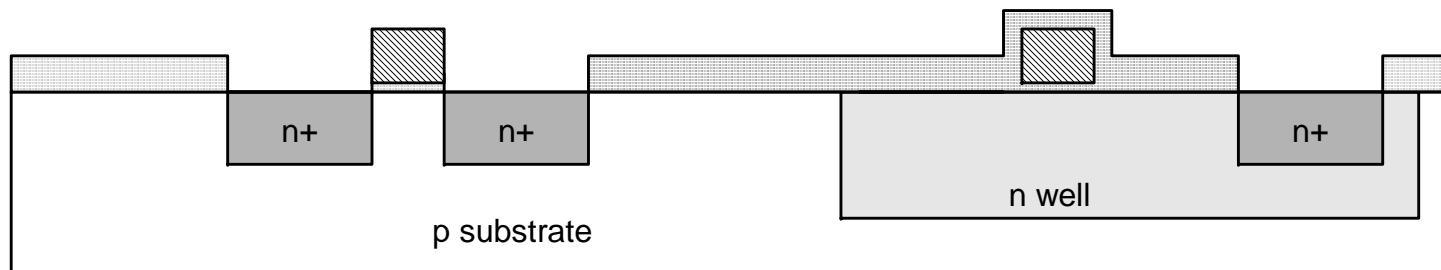
n -Diffusion

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



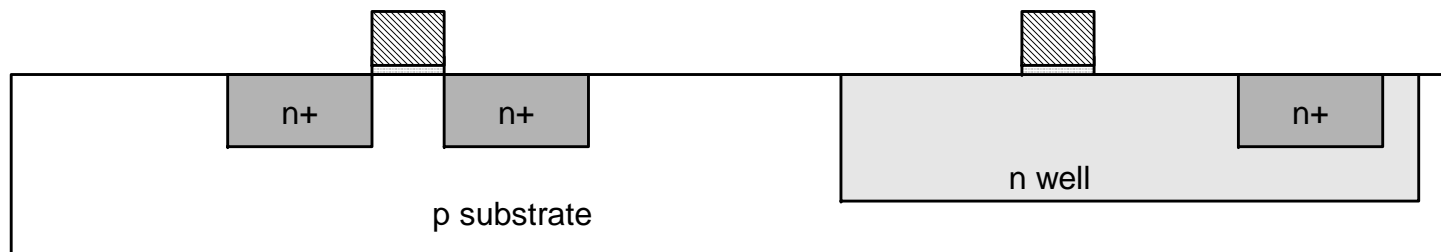
n-Diffusion (cont'd.)

- Historically dopants were diffused
- Usually ion implantation today
- But regions are still called diffusion



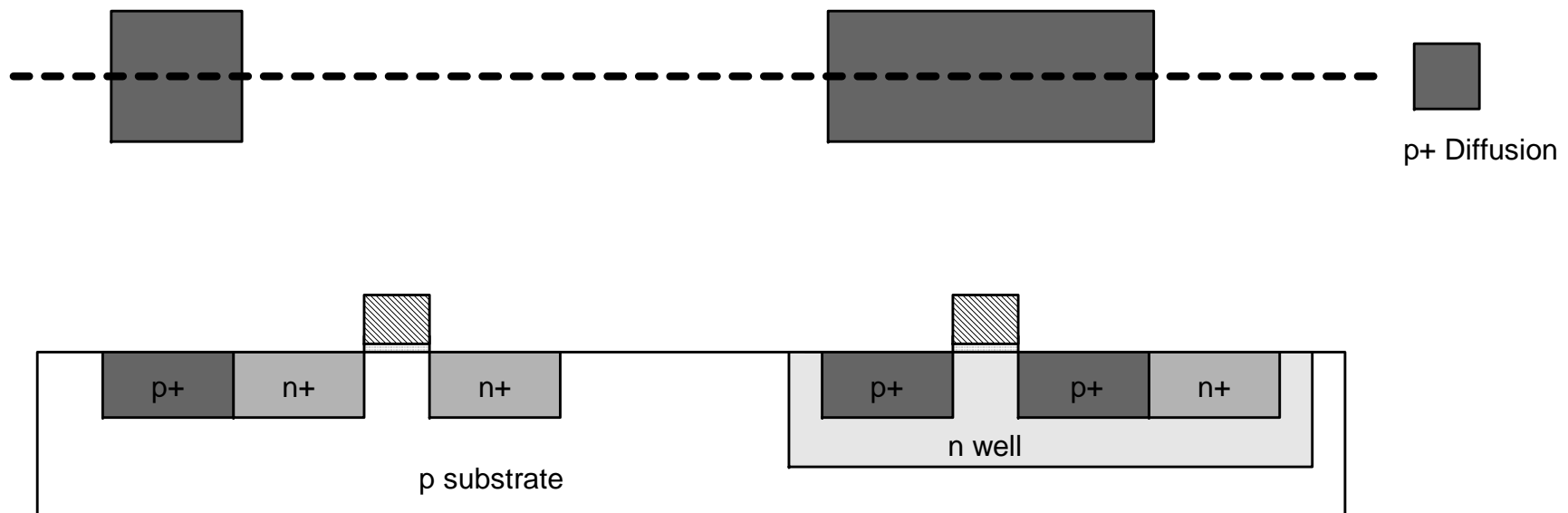
n-Diffusion (cont'd.)

- Strip off oxide to complete patterning step



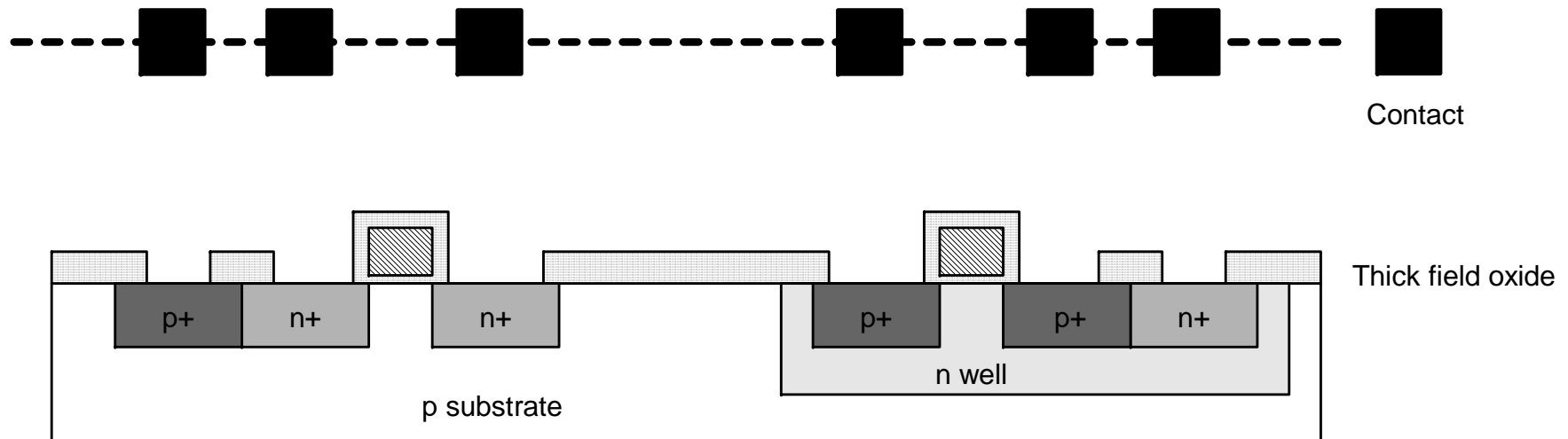
p-Diffusion

- Similar set of steps form p+ diffusion regions for pMOS source and drain and substrate contacts



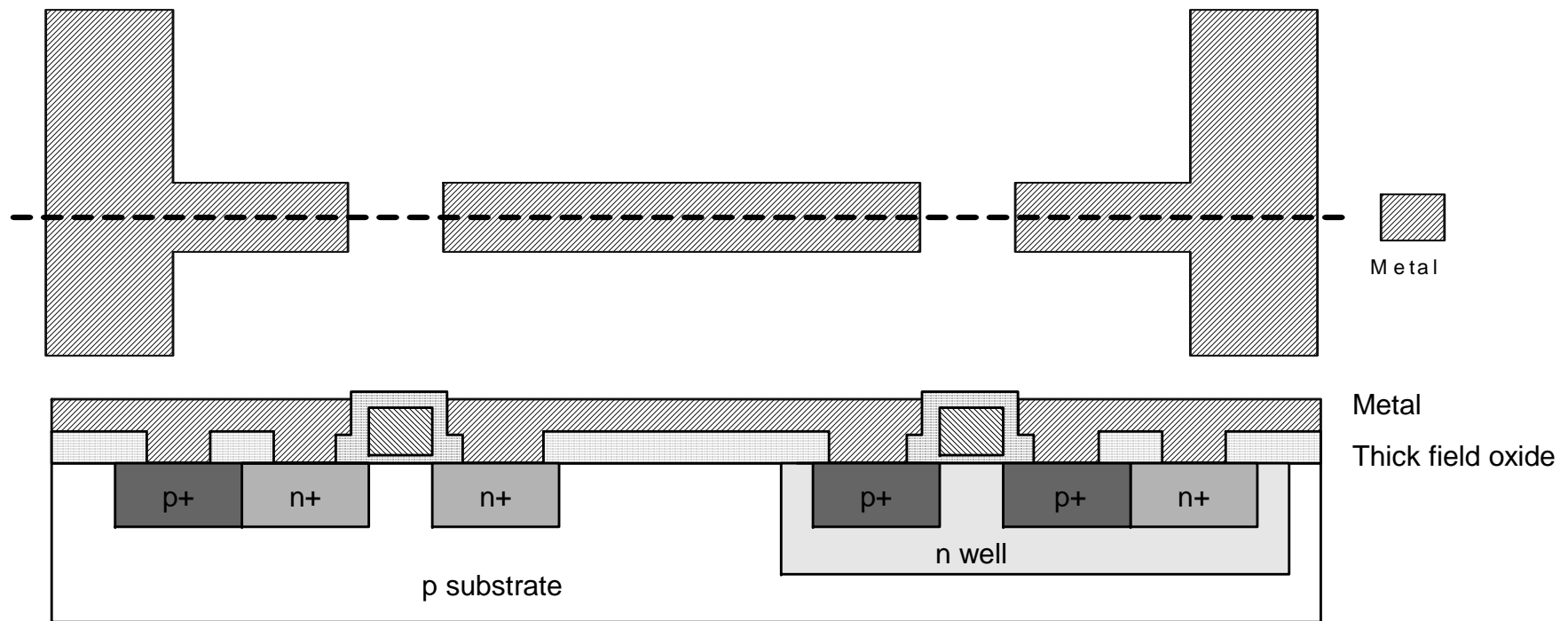
Contacts

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

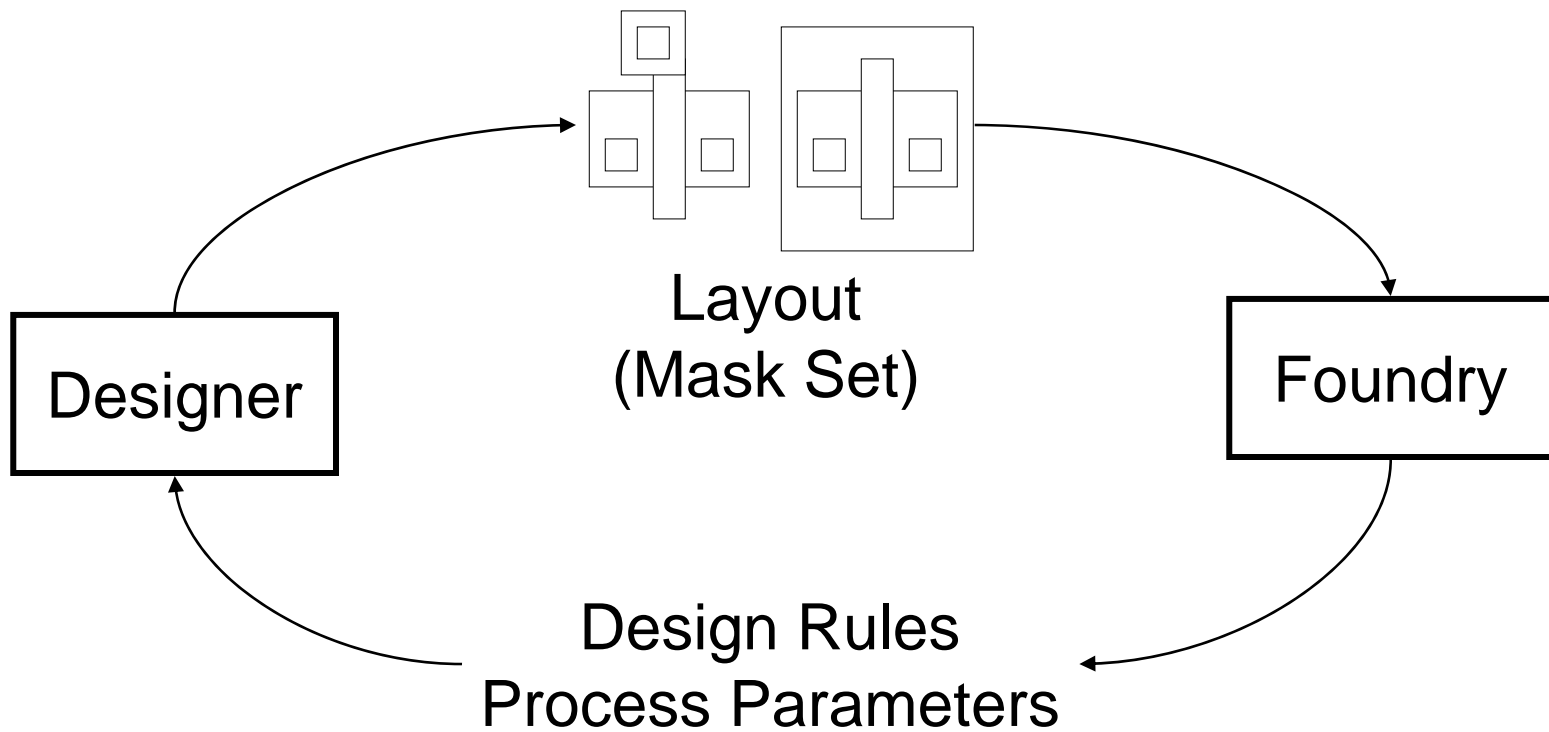


Metallization

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



Foundry Interface



MAGIC MOSIS SCMOS Layers

- 4 types of diffusion
 - Normal (forms transistor)
 - ndiff
 - pdiff
 - Diffusion for well contacts
 - nohmic
 - pohmic
- Poly
- Metal
 - M1
 - M2

Physical Layers

Physical Masks (simplified)

nwell

active area (thin ox)

poly

threshold adjust (n & p)

implant select (n & p)

contact

metal 1

via

metal 2

glass

Fabrication Constraints On Layout

- Resolution constraints
 - Smallest printable feature / smallest spacing that guarantees no short
 - Depends on lithography and processing steps
 - Resolution often depends on smoothness of surface
- Alignment/overlap constraints
 - Need to align layers (like printing color picture)

MOSIS SCMOS Design Rules

- Allow you to send designs to different fabs
- Chips are specified with set of masks
- Minimum dimensions of masks determine transistor size (and hence speed, cost, and power)
- Feature size 2λ = distance between source and drain
 - Set by minimum width of polysilicon
- Feature size improves 30% every 3 years or so
- Normalize for feature size when describing design rules
- Express rules in terms of λ
 - E.g. $\lambda = 0.3\ \mu\text{m}$ in $0.6\ \mu\text{m}$ process
- Conservative
 - Manhattan layout (only 90 degree angles)

Layout Issues

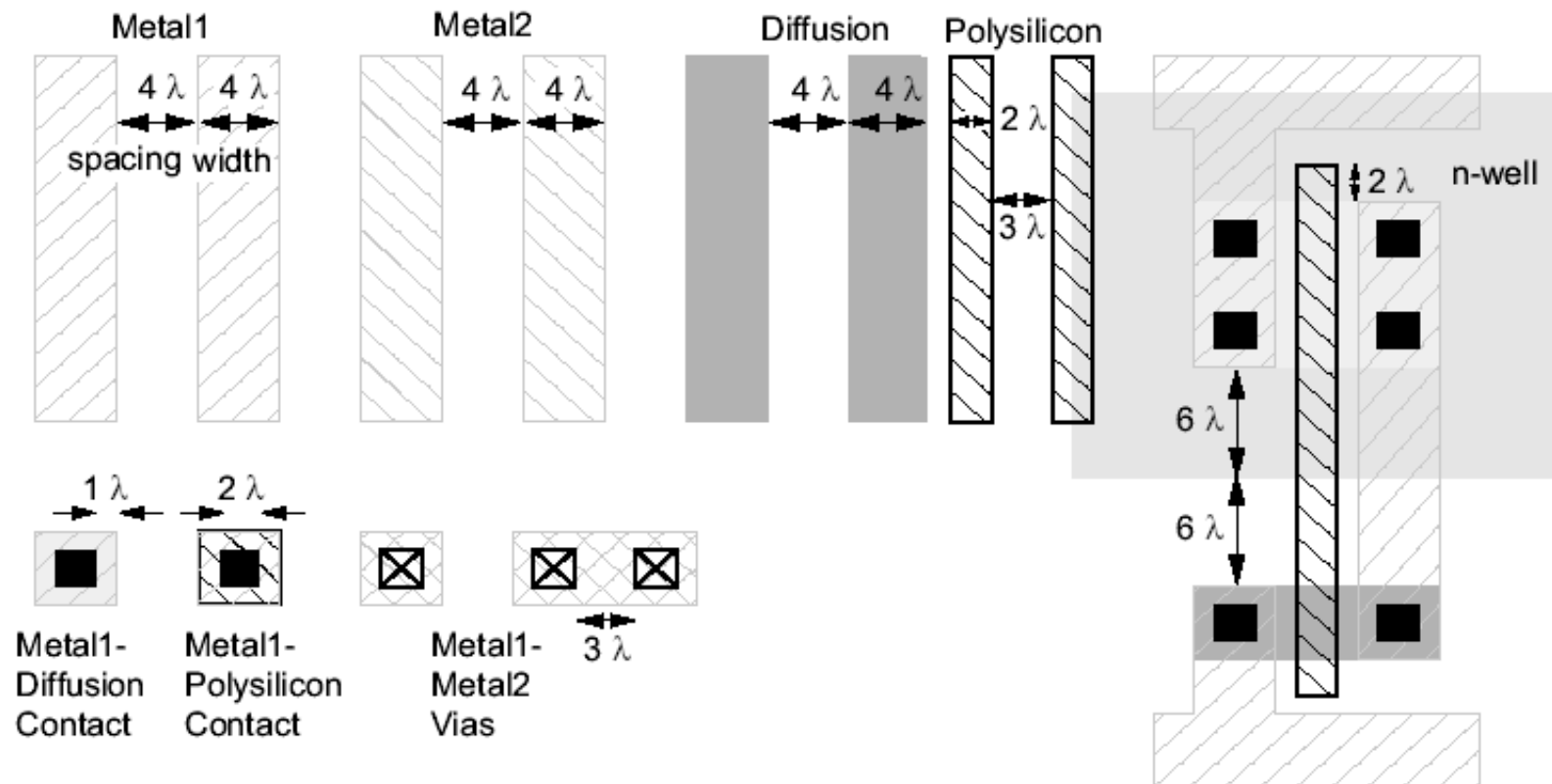
- Two types of diffusion
 - ndiff
 - poly crossing ndiff makes nMOS transistor
 - pdiff
 - poly crossing pdiff makes pMOS transistor
 - Cannot directly connect ndiff and pdiff
 - must connect ndiff to metal and metal to pdiff
 - Cannot get ndiff too close to pdiff because of wells
 - large spacing rule between ndiff and pdiff
 - need to group nMOS transistors together and pMOS transistors together

Basic Layout Planning

- Need to route power and ground (in metal)
- Keep nMOS devices near nMOS devices and pMOS devices near pMOS devices
 - nMOS near ground and pMOS near Vdd
- Run poly vertically and diffusion horizontally with m1 horizontally
- Keep diffusion wires as short as possible
 - just enough to make transistors
- All long wires in m1 and m2

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, $0.6 \mu\text{m}$ long

