EE4800 CMOS Digital IC Design & Analysis

Lecture 1 Introduction

Zhuo Feng

Prof. Zhuo Feng

- Office: EERC 730
- Phone: 487-3116
- Email: zhuofeng@mtu.edu

Class Website

- **http://www.ece.mtu.edu/~zhuofeng/EE4800Fall2010.html**
- Check the class website for lecture materials, assignments and announcements

Schedule

- TR 12:35pm-13:50pm EERC 227
- Office hours: TR 4:30pm – 5:30pm or by appointments

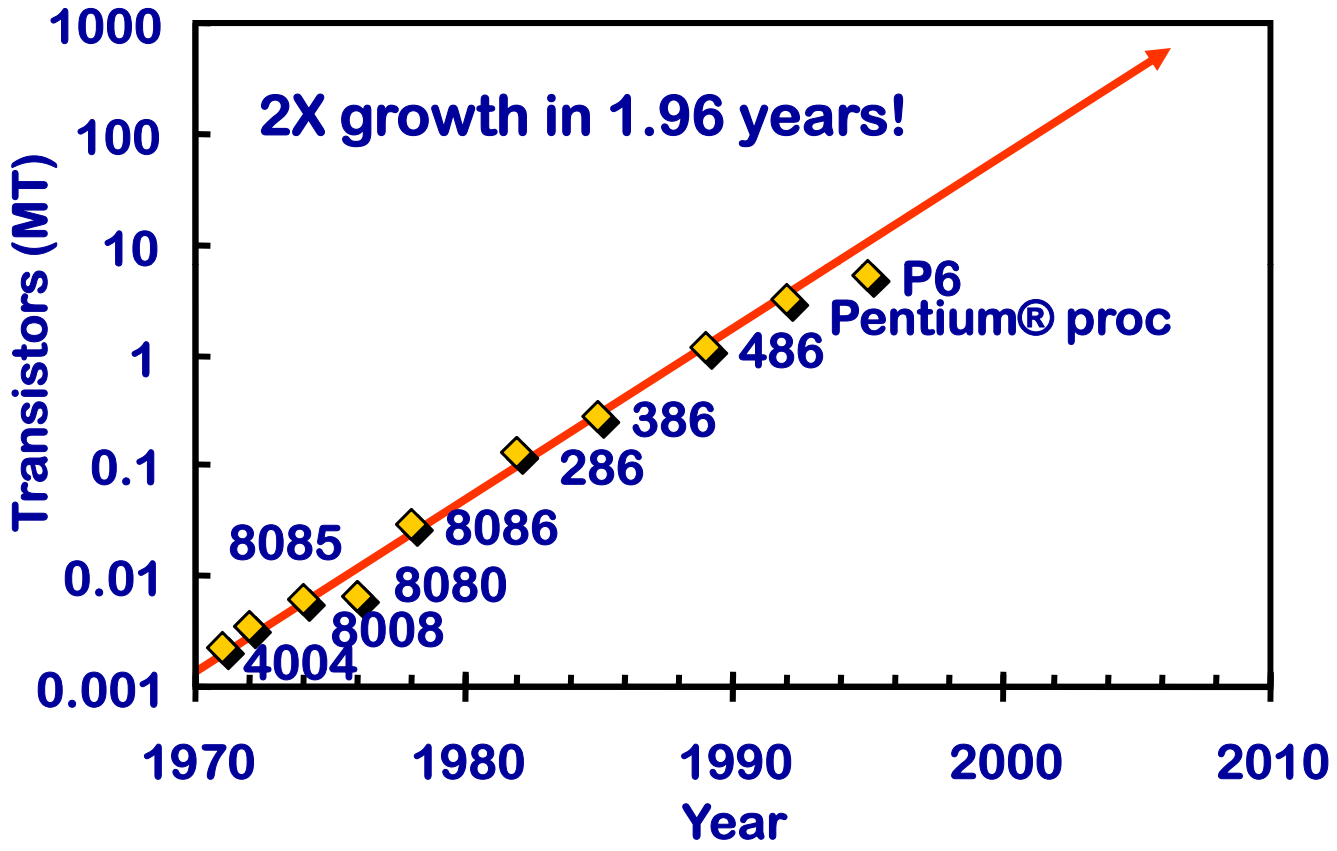
Topics (tentative)

- CMOS circuit overview
- Fabrication and layout
- MOS Transistor I-V Characteristics
- DC & Transient Responses
- Delay and Power Estimation
- Logic Effort
- Interconnect
- Combinational Circuits
- Sequential Circuits
- Clock Distribution
- Memory
- Package, Power, I/O

Grading Policy

- Homework: 40%
- Quizzes 10%
- Mid-term Exam: 20%
- Final Exam: 30%
- Late homework: 50% penalty/day.
- Letter Grades:
  - A: 85~100; AB: 80~84; B: 75~79; BC: 70~74; C: 65~69; D: 60~64; F: 0~59

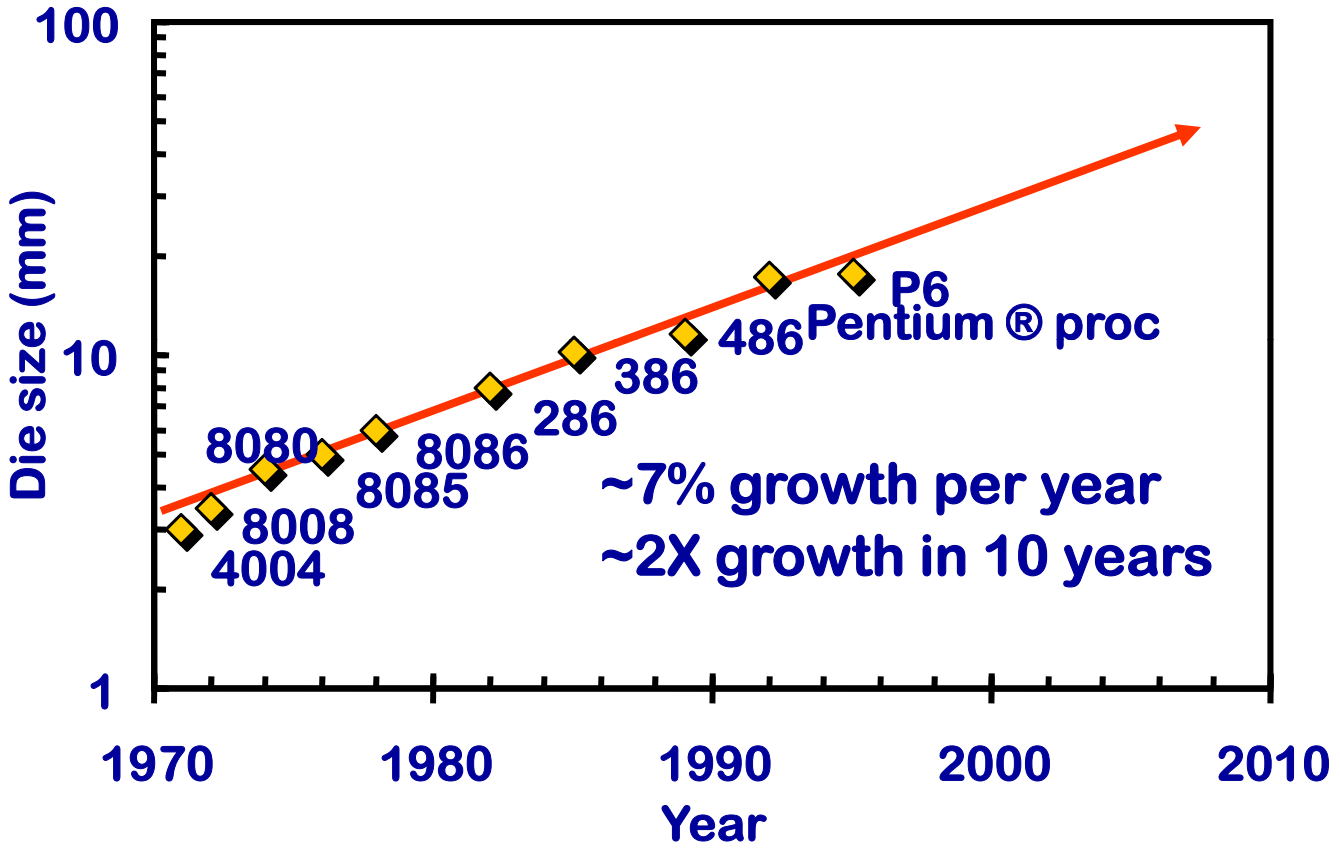
Moore's law in Microprocessors



Transistors on Lead Microprocessors double every 2 years

Courtesy, Intel

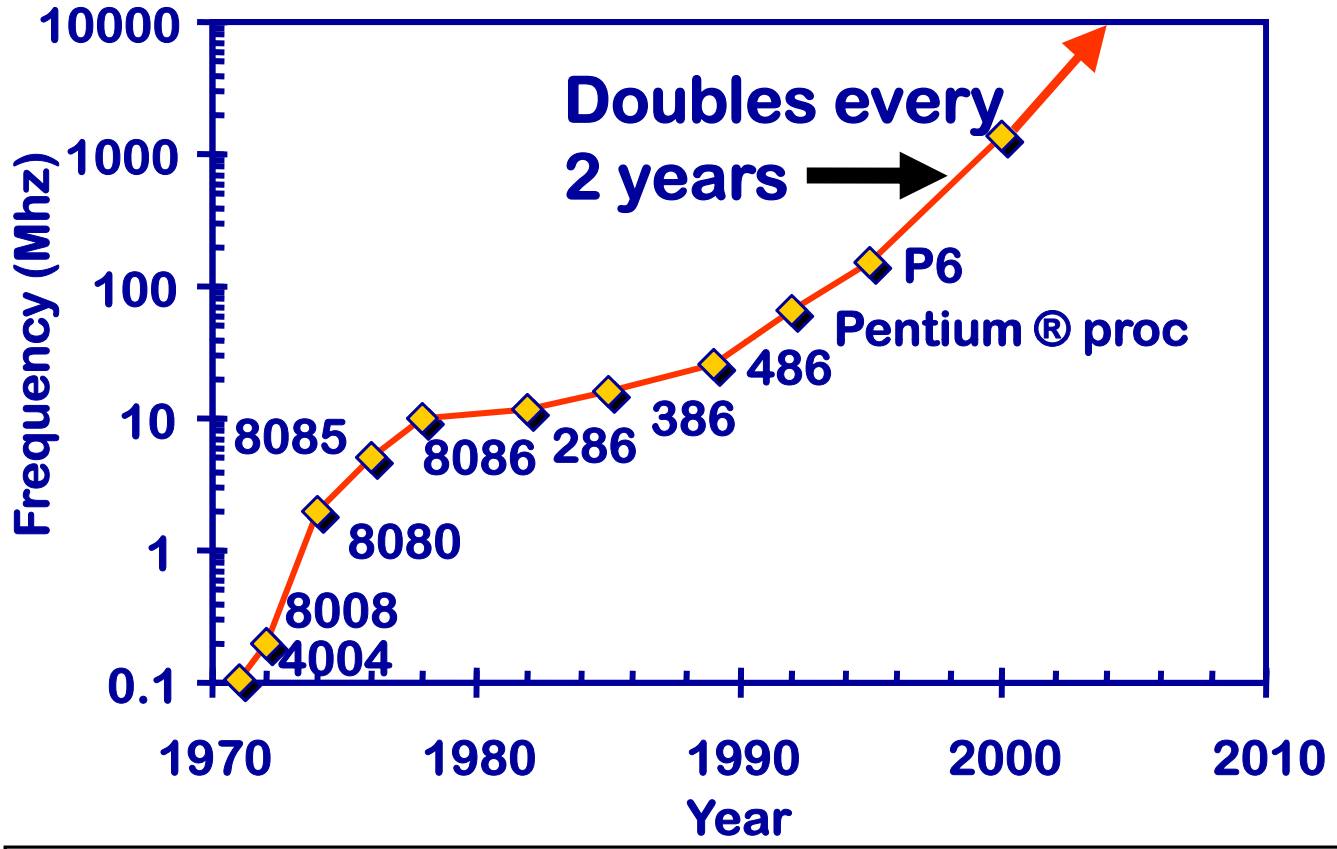
Die Size Growth



Die size grows by 14% to satisfy Moore's Law

Courtesy, Intel

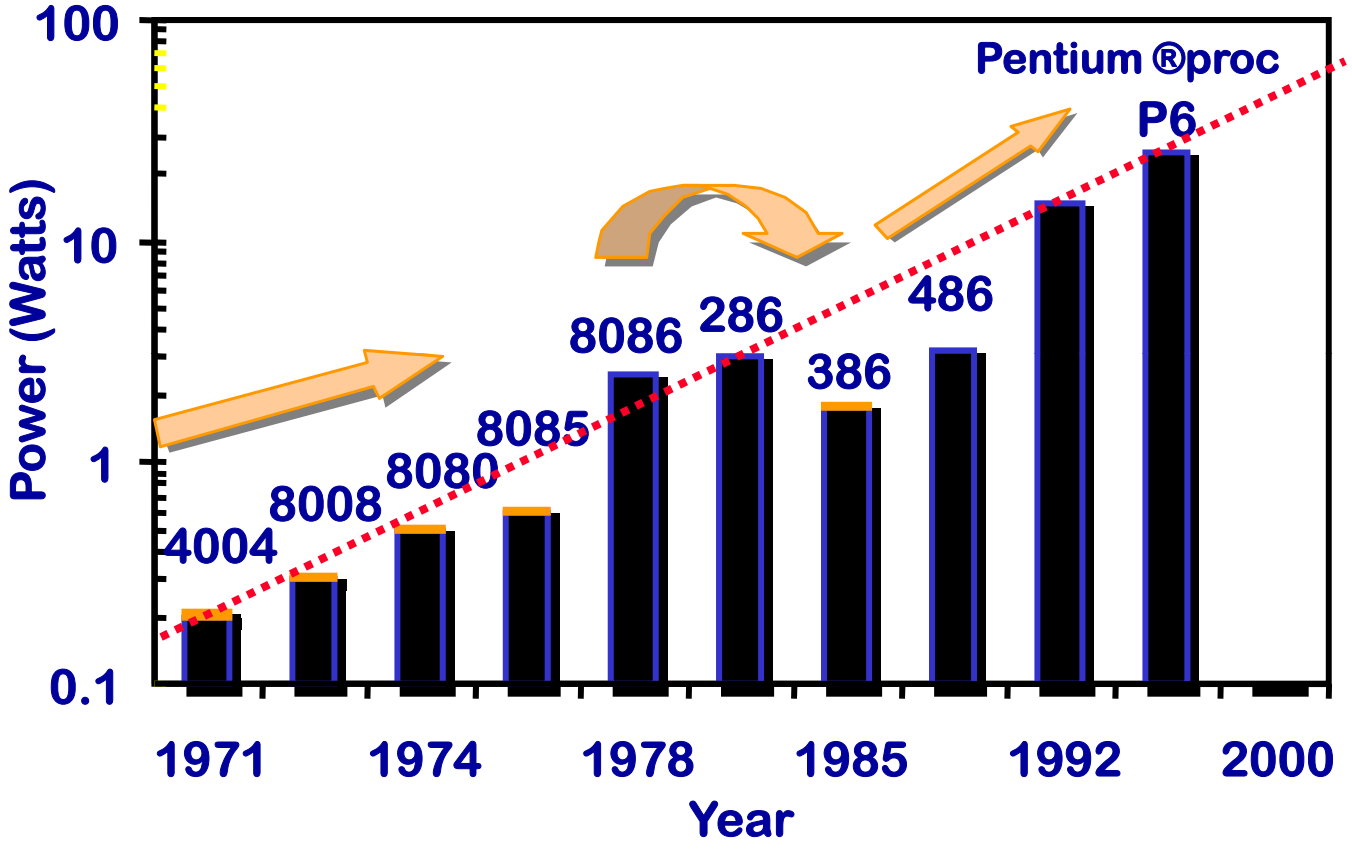
Frequency



Lead Microprocessors frequency doubles every 2 years

Courtesy, Intel

Power Dissipation



Lead Microprocessors power continues to increase

Courtesy, Intel

Why Scaling?

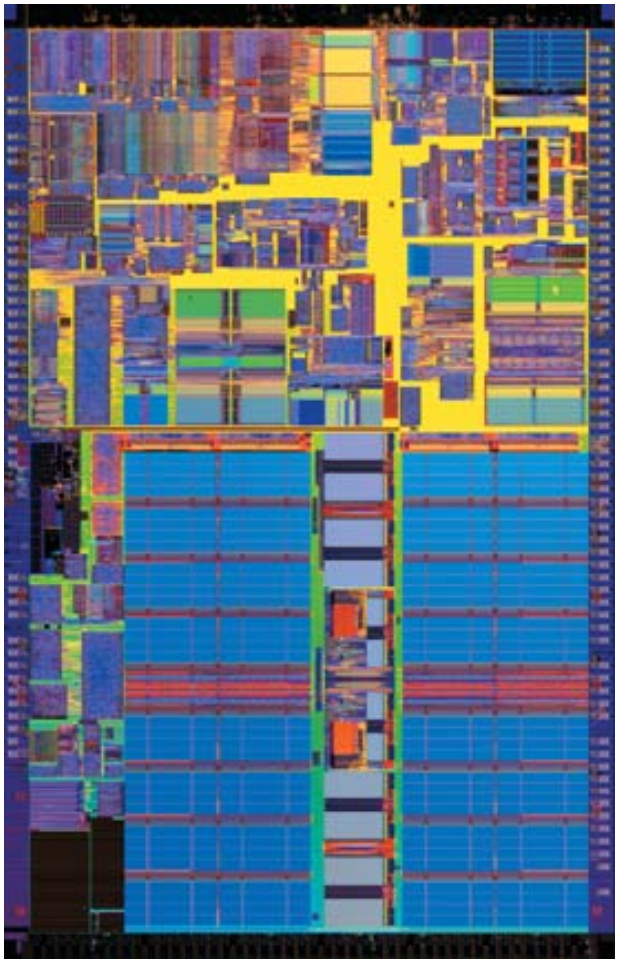
- **Technology shrinks by ~0.7 per generation**
- **With every generation can integrate 2x more functions on a chip; chip cost does not increase significantly**
- **Cost of a function decreases by 2x**
- **But …**
  - How to design chips with more and more functions?
  - Design engineering population does not double every two years…
- **Hence, a need for more efficient design methods**
  - Exploit different levels of abstraction

Pentium 4

- **Deep pipeline (2001)**
  - Very fast clock
  - 256-1024 KB L2$
- **Characteristics**
  - 180 – 65 nm process
  - 42-125M transistors
  - 1.4-3.4 GHz
  - Up to 160 W
  - 32/64-bit word size
  - 478-pin PGA
- **Units start to become invisible on this scale**

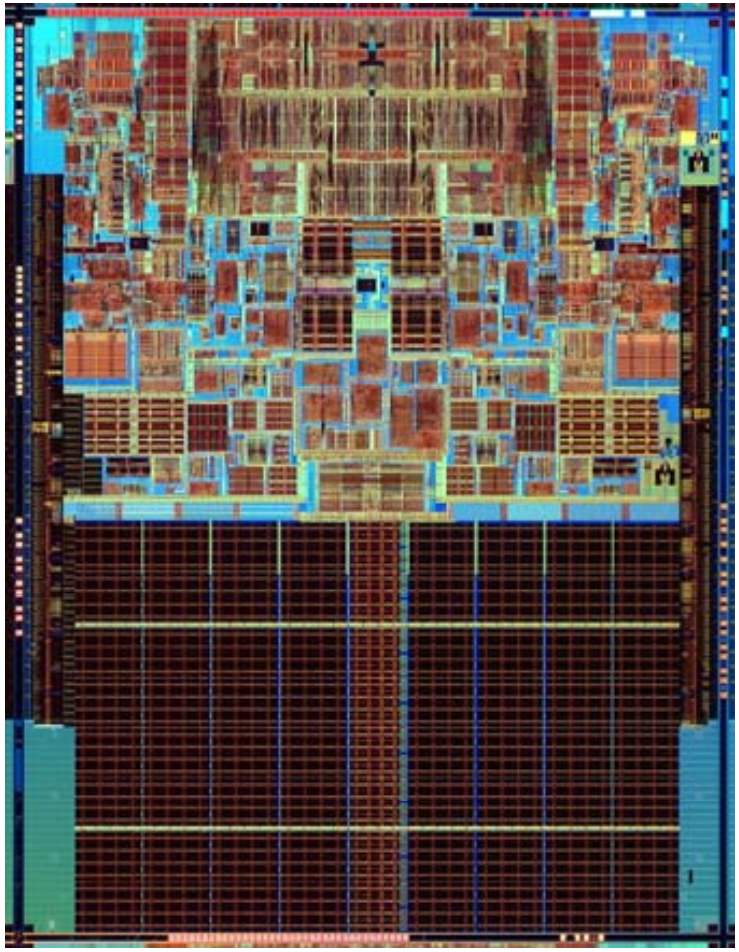
Pentium M

- **Pentium III derivative**
  - Better power efficiency
  - 1-2 MB L2$
- **Characteristics**
  - 130 – 90 nm process
  - 140M transistors
  - 0.9-2.3 GHz
  - 6-25 W
  - 32-bit word size
  - 478-pin PGA
- **Cache dominates chip area**



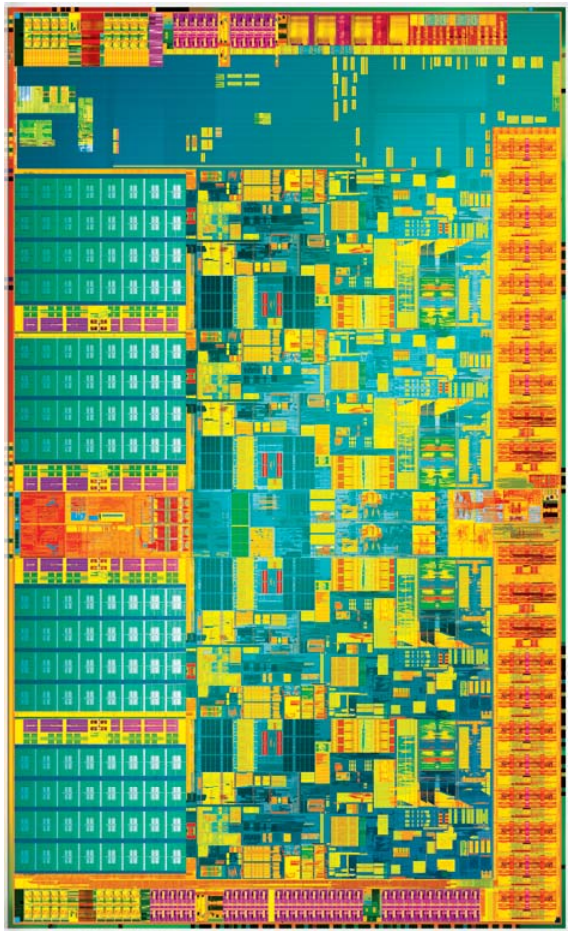
Core2 Duo

- **Dual core (2006)**
  - 1-2 MB L2$ / core
- **Characteristics**
  - 65-45 nm process
  - 291M transistors
  - 1.6-3+ GHz
  - 65 W
  - 32/64 bit word size
  - 775 pin LGA
- **Much better performance/power efficiency**



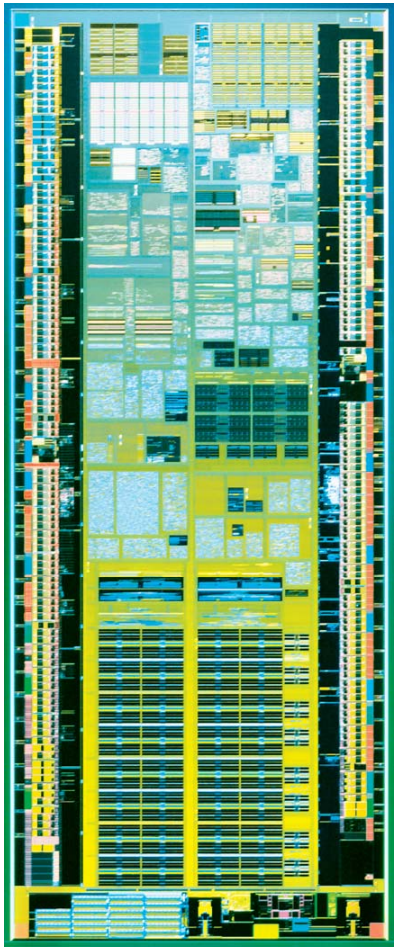
Core i7

- **Quad core (& more)**
  - Pentium-style architecture
  - 2 MB L3$ / core
- **Characteristics**
  - 45-32 nm process
  - 731M transistors
  - 2.66-3.33+ GHz
  - Up to 130 W
  - 32/64 bit word size
  - 1366-pin LGA
  - Multithreading
- **On-die memory controller**

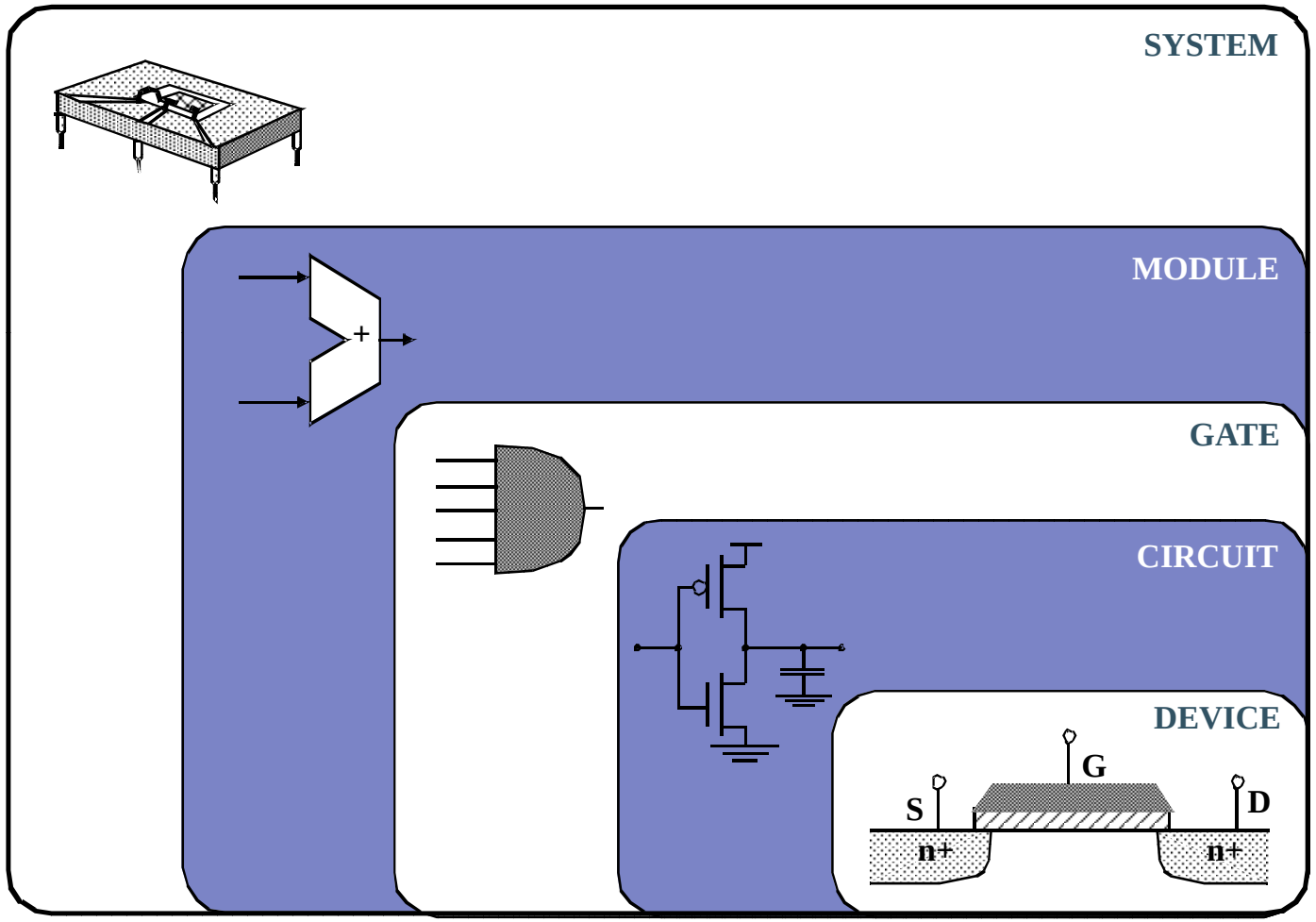


Atom

- **Low power CPU for netbooks**
  - Pentium-style architecture
  - 512KB+ L2$
- **Characteristics**
  - 45-32 nm process
  - 47M transistors
  - 0.8-1.8+ GHz
  - 1.4-13 W
  - 32/64-bit word size
  - 441-pin FCBGA
- **Low voltage (0.7 – 1.1 V) operation**
  - Excellent performance/power



Design Abstraction Levels



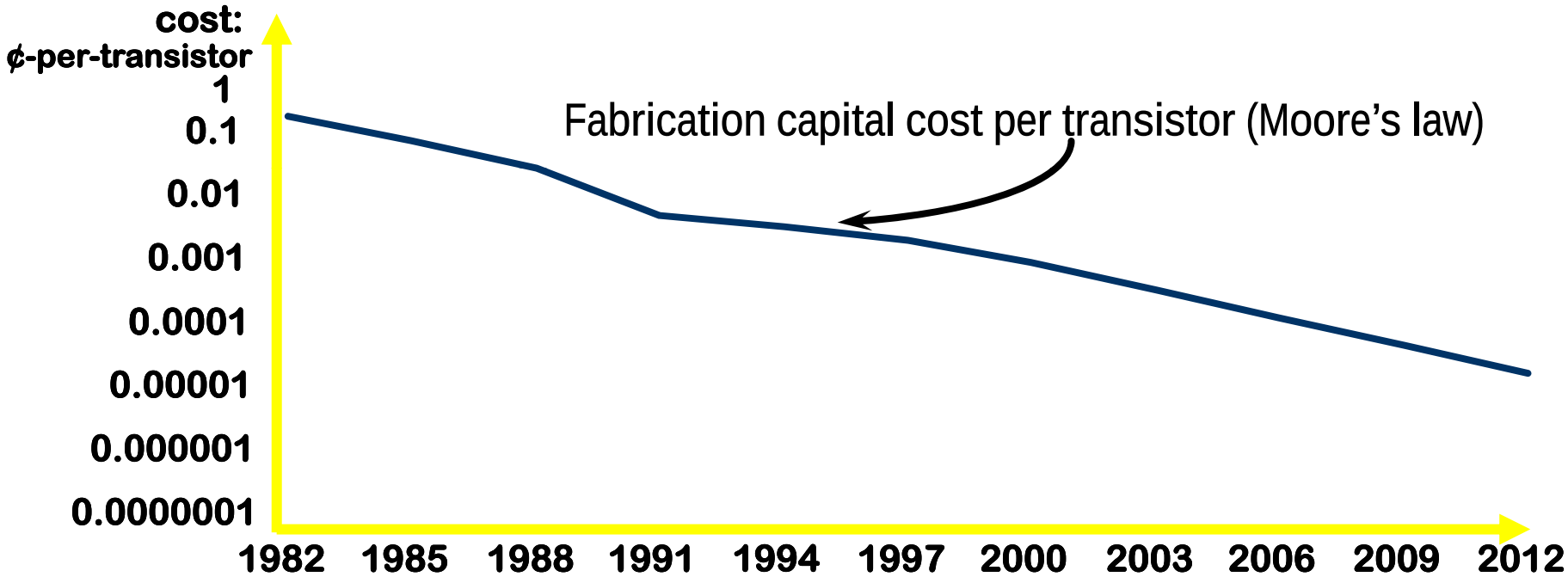
Design Metrics

- **How to evaluate performance of a digital circuit (gate, block, …)?**
  - Cost
  - Reliability
  - Scalability
  - Speed (delay, operating frequency)
  - Power dissipation
  - Energy to perform a function

Cost of Integrated Circuits

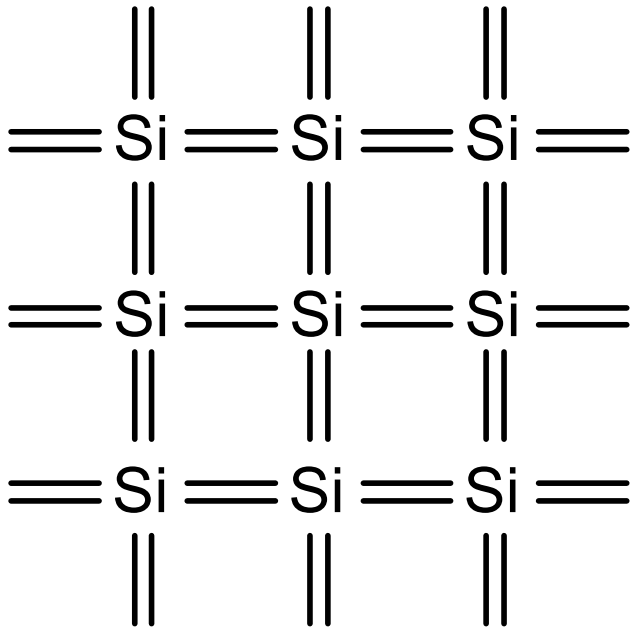
- **NRE (non-recurrent engineering) costs**
  - ▶ design time and effort, mask generation
  - ▶ one-time cost factor
- **Recurrent costs**
  - ▶ silicon processing, packaging, test
  - ▶ proportional to volume
  - ▶ proportional to chip area

Cost per Transistor



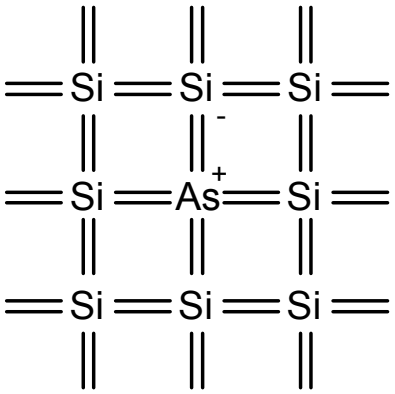
Silicon Lattice

- Transistors are built on a silicon substrate
- Silicon is a Group IV material
- Forms crystal lattice with bonds to four neighbors

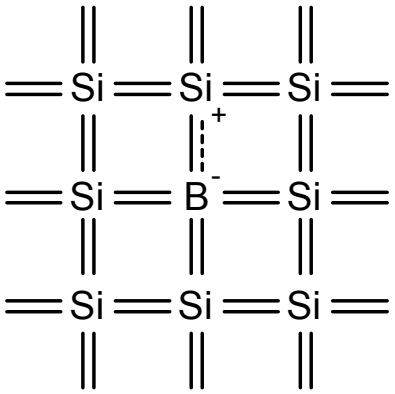


Dopants

- Silicon is a semiconductor
- Pure silicon has no free carriers and conducts poorly
- Adding dopants increases the conductivity
- Group V: extra electron (n-type)
- Group III: missing electron, called hole (p-type)



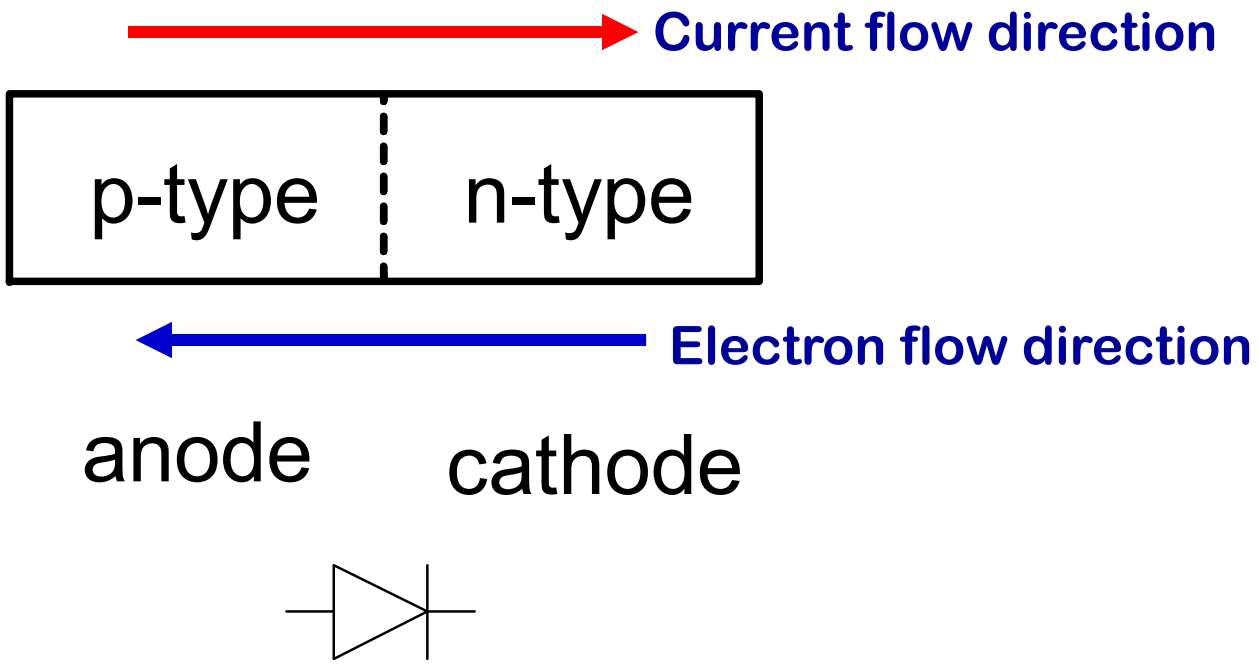
N-type



P-type

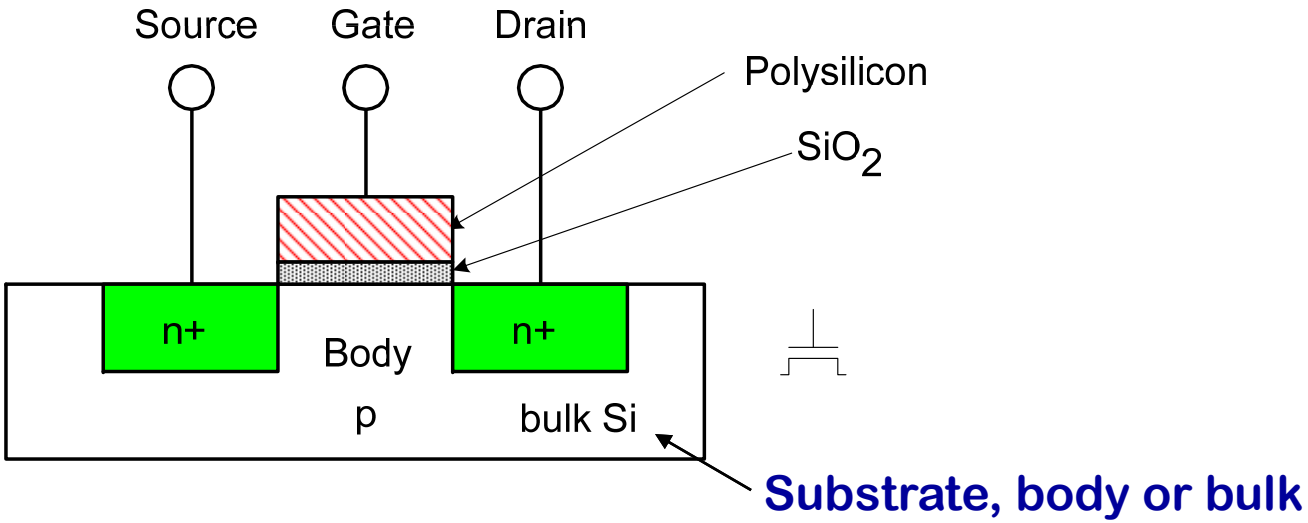
P-N Junctions

- A junction between p-type and n-type semiconductor forms a diode.
- Current flows only in one direction



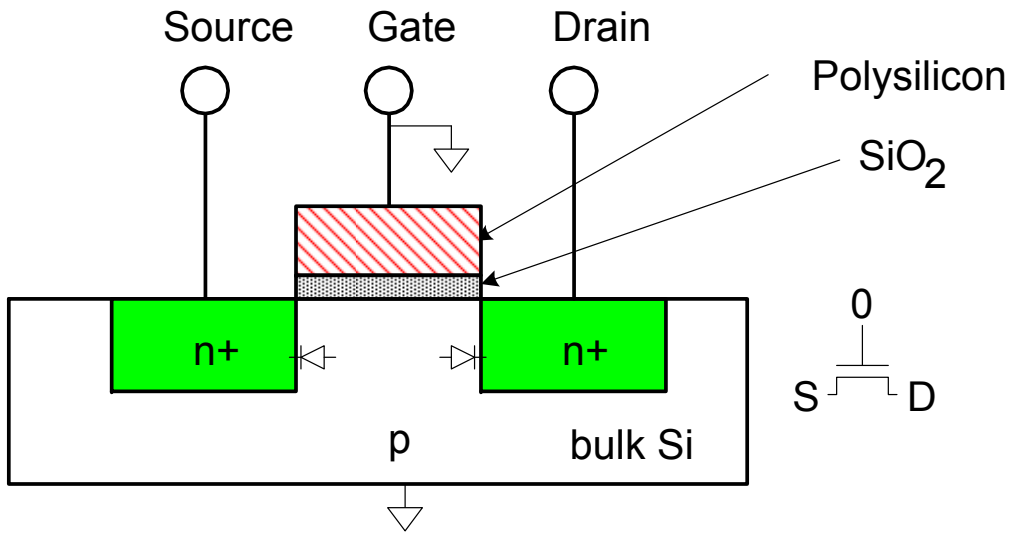
NMOS Transistor

- **Four terminals: gate, source, drain, body**
- **Gate – oxide – body stack looks like a capacitor**
  - Gate and body are conductors
  - SiO_2 (oxide) is a very good insulator
  - Called metal – oxide – semiconductor (MOS) capacitor
  - Even though gate is no longer made of metal



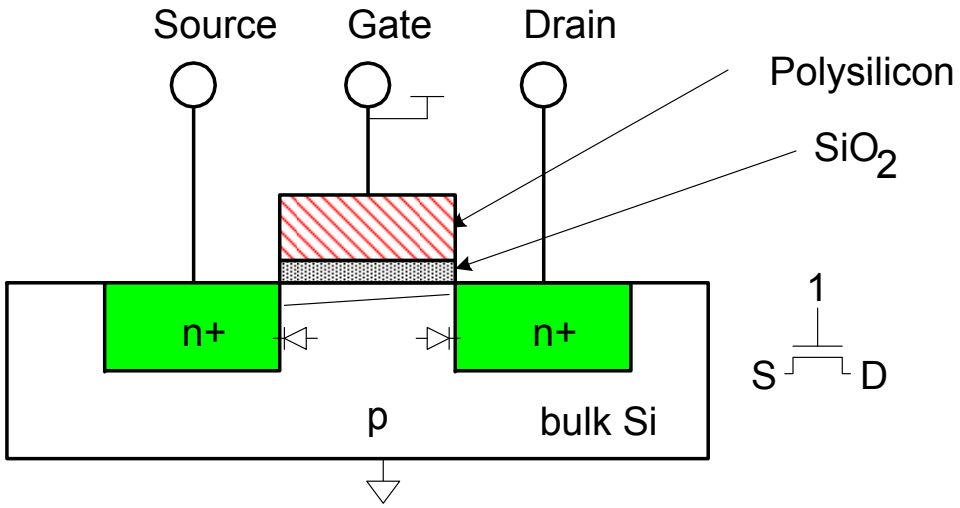
NMOS Operation

- **Body is commonly tied to ground (0 V)**
- **When the gate is at a low voltage:**
  - P-type body is at low voltage
  - Source-body and drain-body diodes are OFF
  - No current flows, transistor is OFF



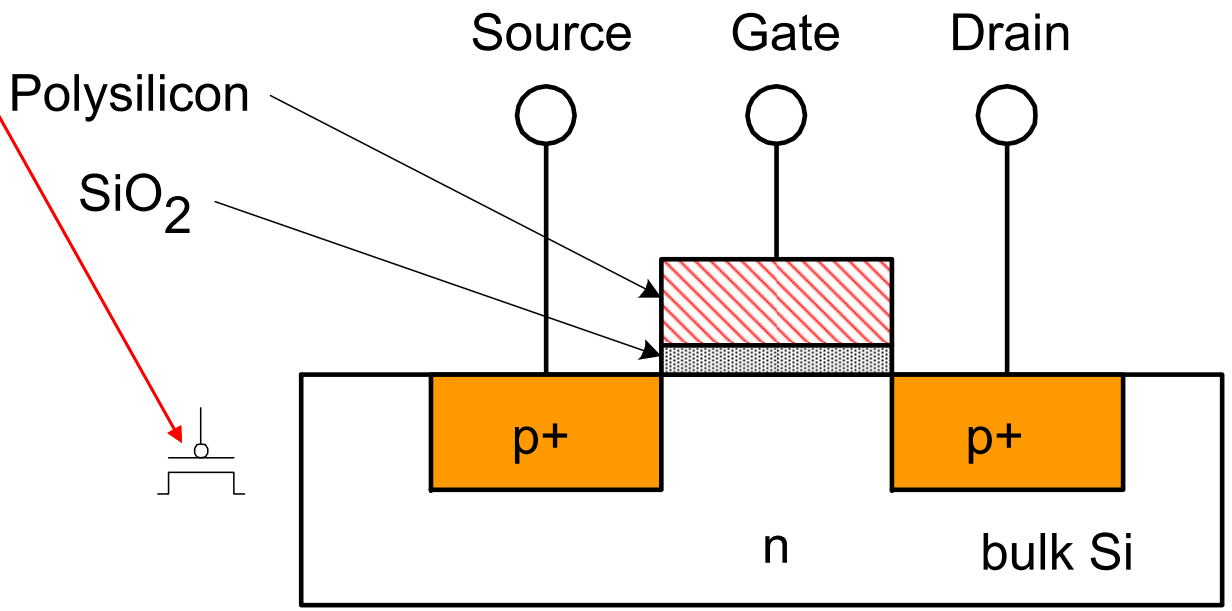
NMOS Operation Cont.

- **When the gate is at a high voltage:**
  - Positive charge on gate of MOS capacitor
  - Negative charge attracted to body
  - Inverts a channel under gate to n-type
  - Now current can flow through n-type silicon from source through channel to drain, transistor is ON



PMOS Transistor

- **Similar, but doping and voltages reversed**
  - Body tied to high voltage (V_{DD})
  - Gate low: transistor ON
  - Gate high: transistor OFF
  - Bubble indicates inverted behavior

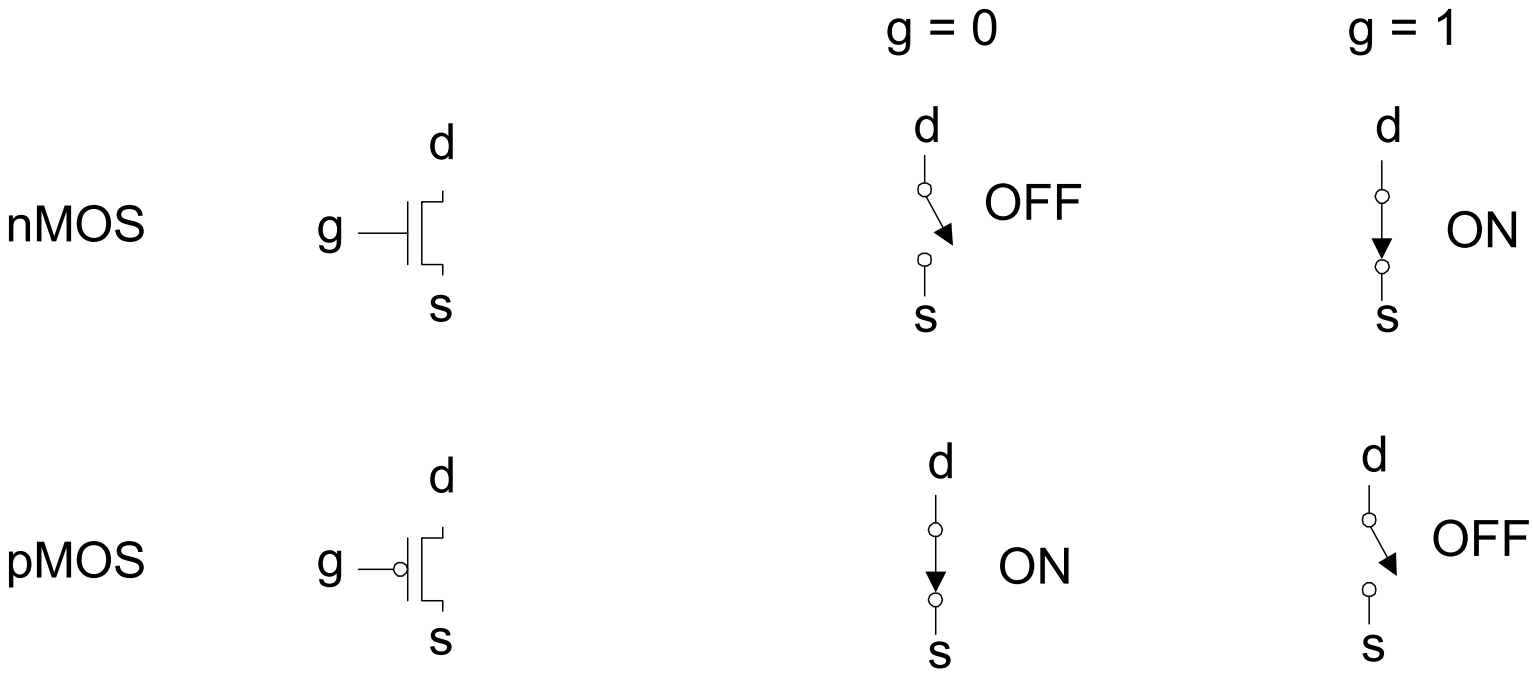


Power Supply Voltage

- **GND = 0 V**
- **In 1980's, V_{DD} = 5V**
- **V_{DD} has decreased in modern processes**
  - ► High V_{DD} would damage modern tiny transistors
  - ► Lower V_{DD} saves power
- **V_{DD} = 3.3, 2.5, 1.8, 1.5, 1.2, 1.0, …**

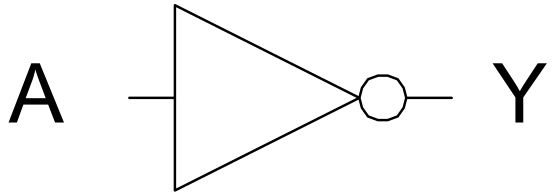
Transistors as Switches

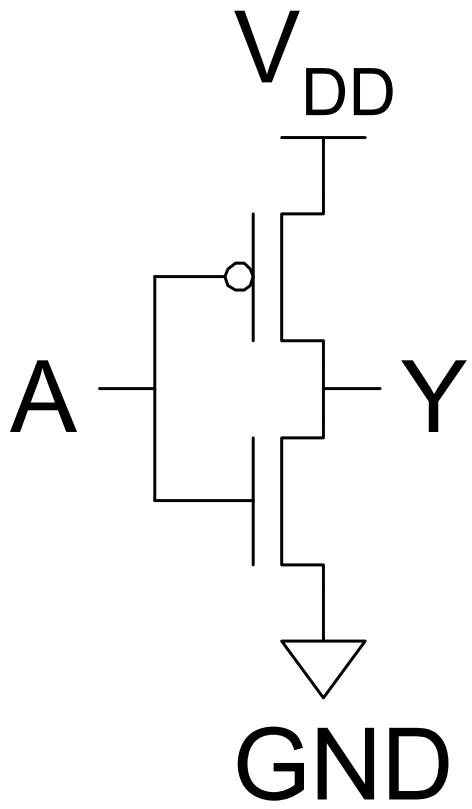
- We can view MOS transistors as electrically controlled switches
- Voltage at gate controls path from source to drain



CMOS Inverter

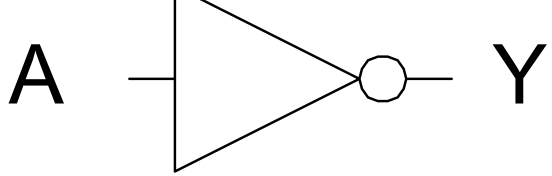
A	Y
0	
1	

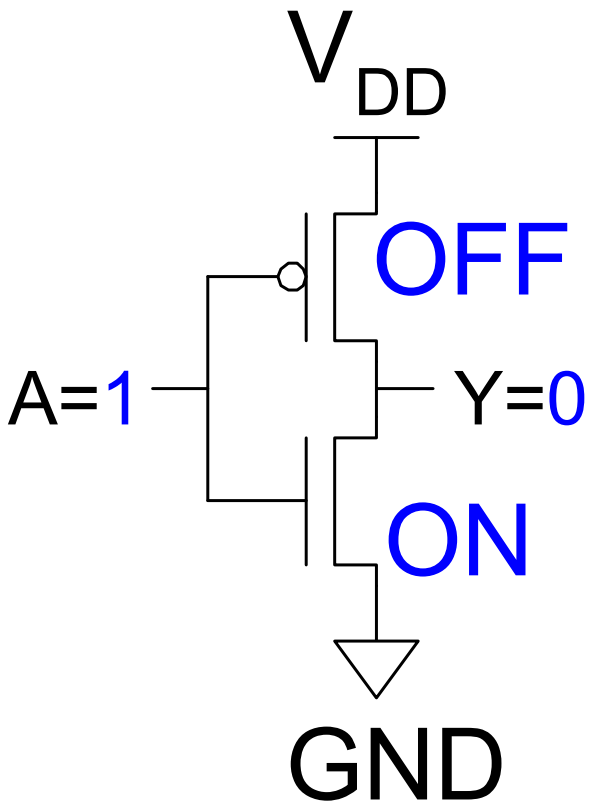




CMOS Inverter

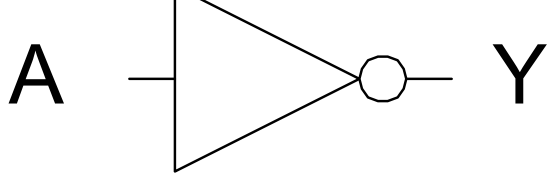
A	Y
0	
1	0

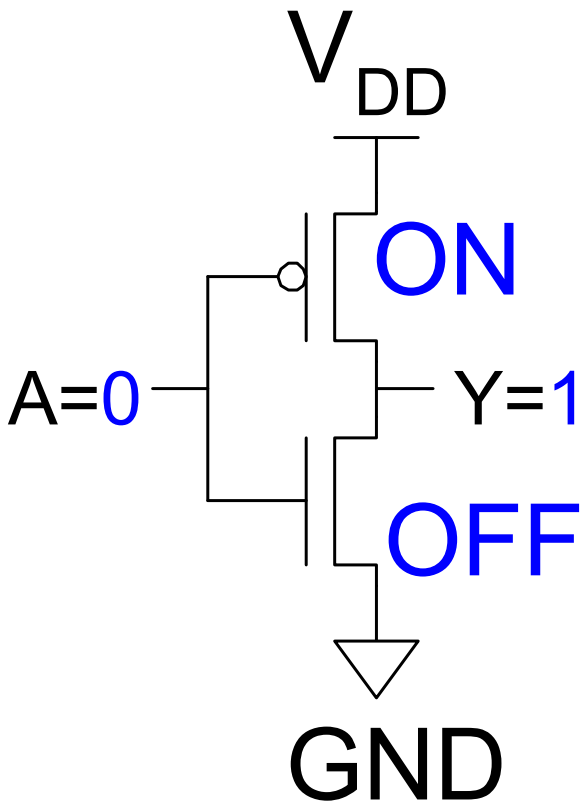




CMOS Inverter

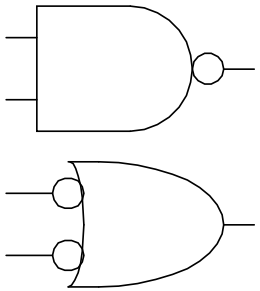
A	Y
0	**1**
1	0

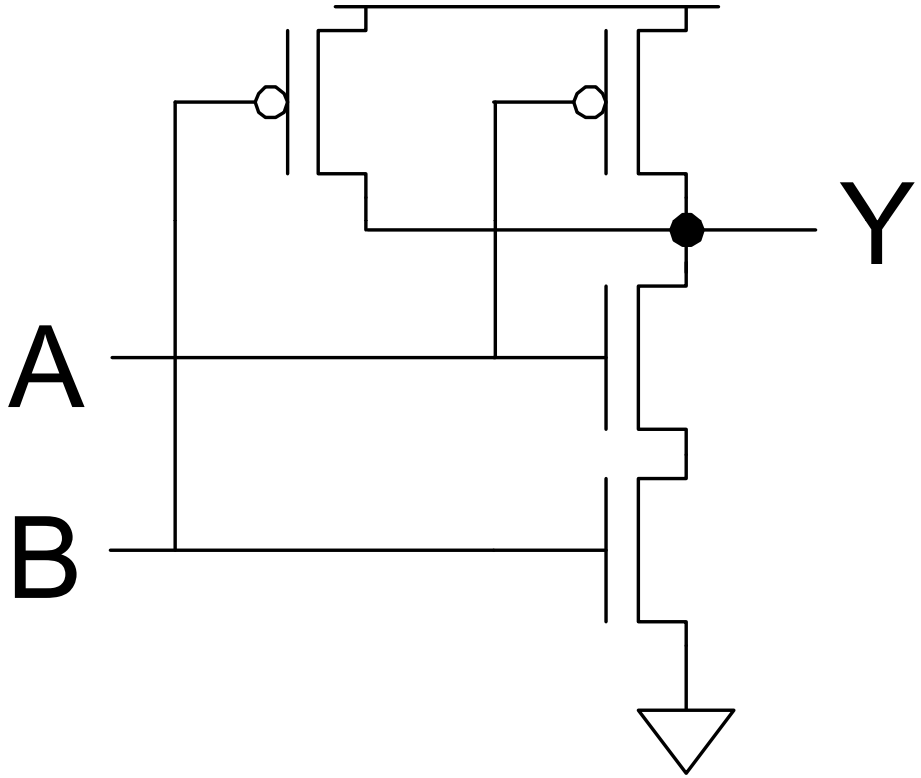




CMOS NAND Gate

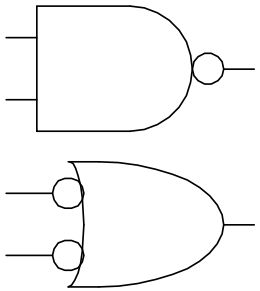
A	B	Y
0	0	
0	1	
1	0	
1	1	

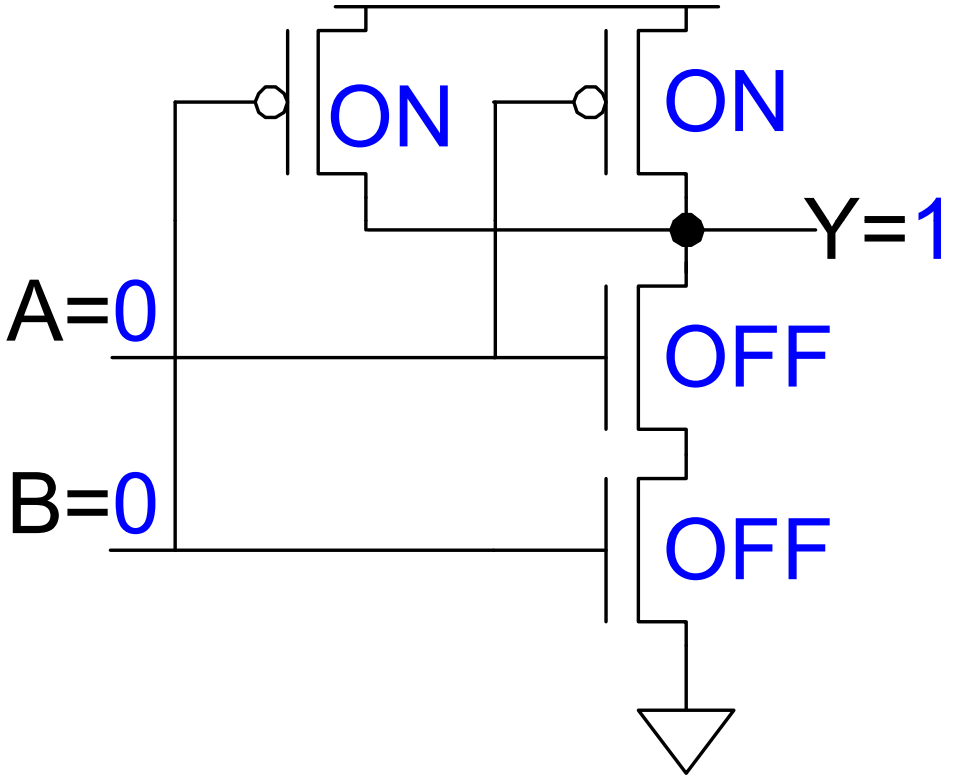




CMOS NAND Gate

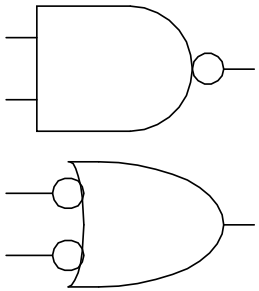
A	B	Y
0	**0**	**1**
0	1	
1	0	
1	1	

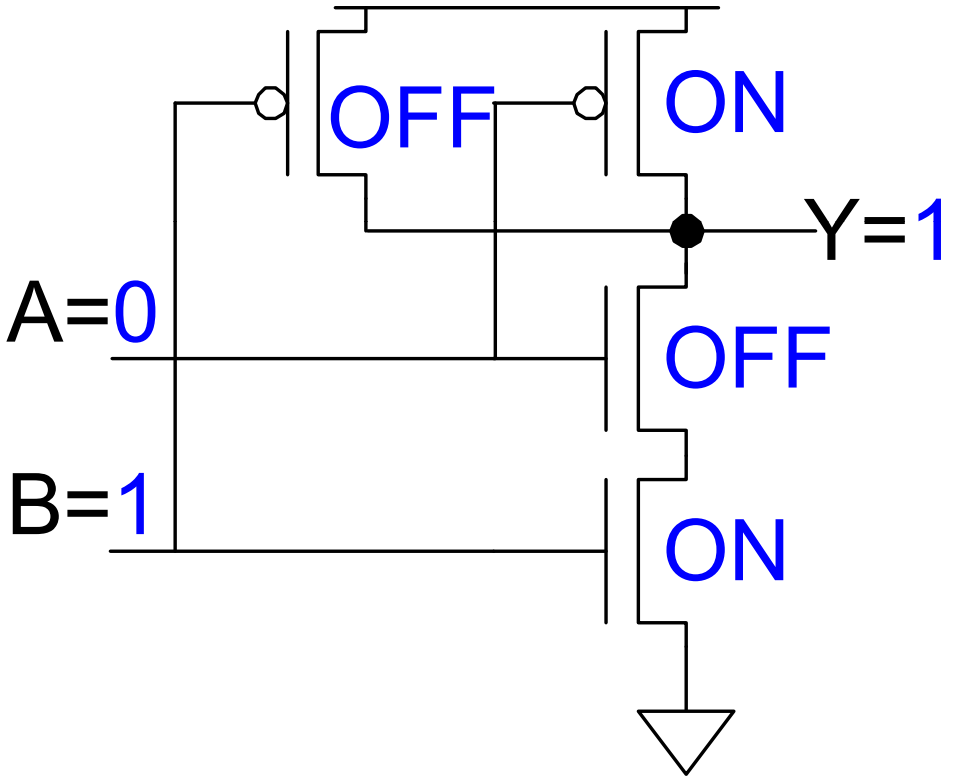




CMOS NAND Gate

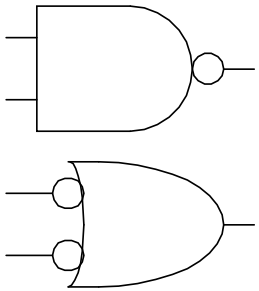
A	B	Y
0	0	1
0	**1**	**1**
1	0	
1	1	

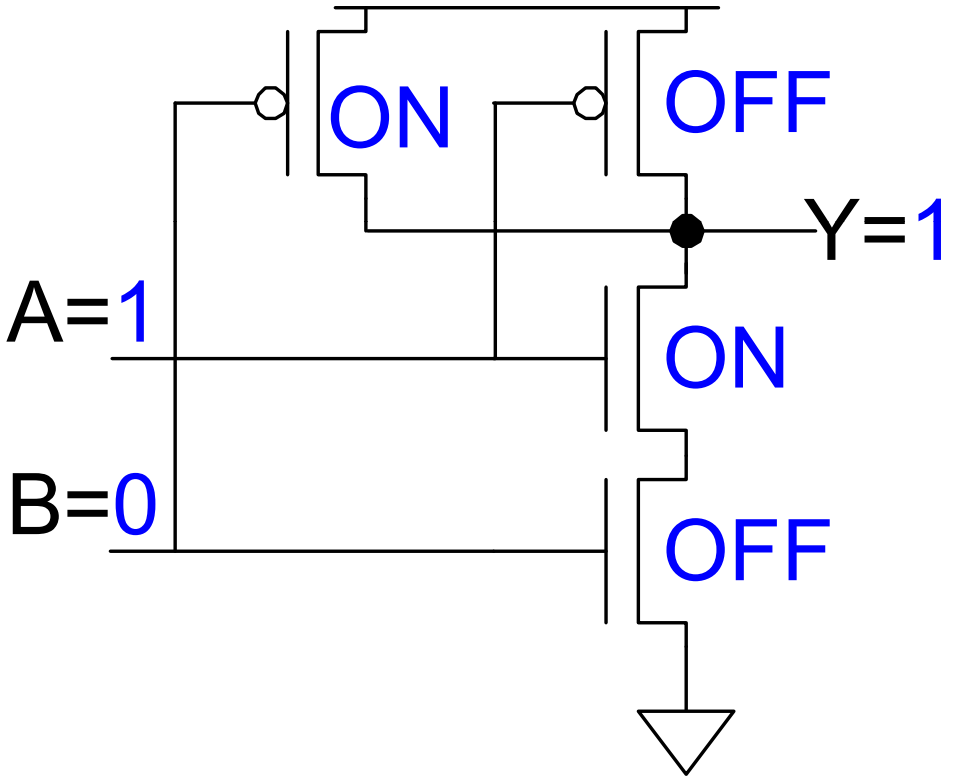




CMOS NAND Gate

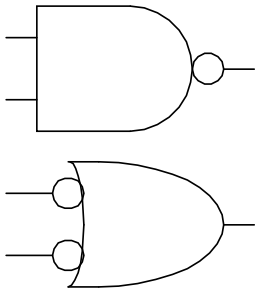
A	B	Y
0	0	1
0	1	1
1	**0**	**1**
1	1	

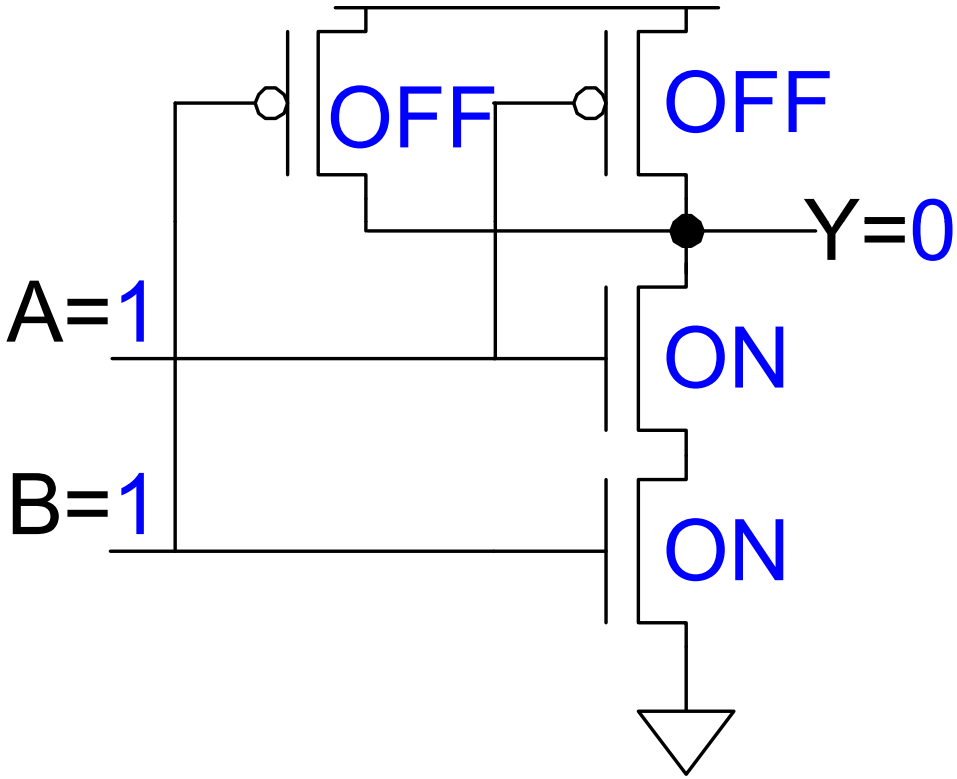




CMOS NAND Gate

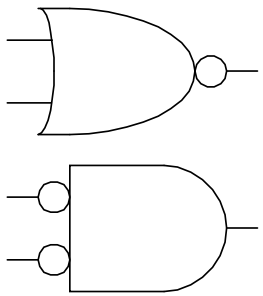
A	B	Y
0	0	1
0	1	1
1	0	1
1	**1**	**0**

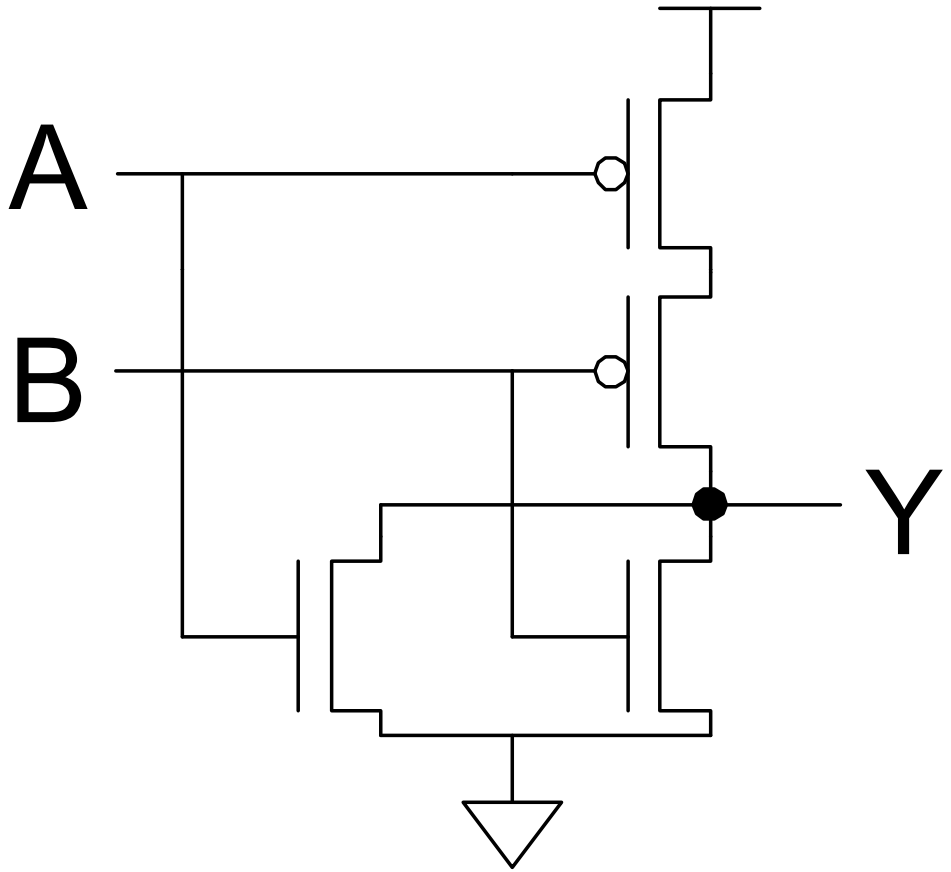




CMOS NOR Gate

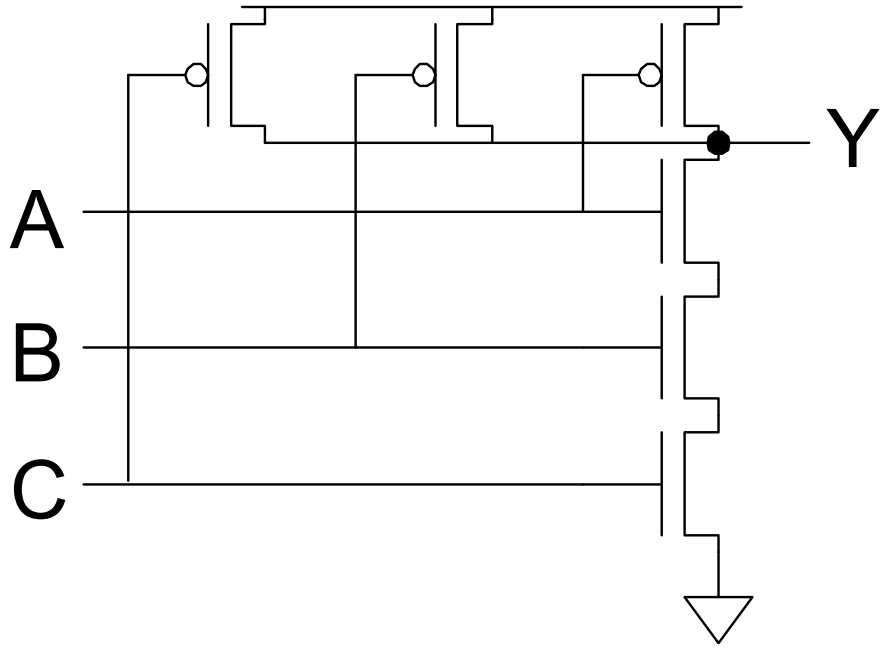
A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0





3-input NAND Gate

- Y pulls low if ALL inputs are 1
- Y pulls high if ANY input is 0

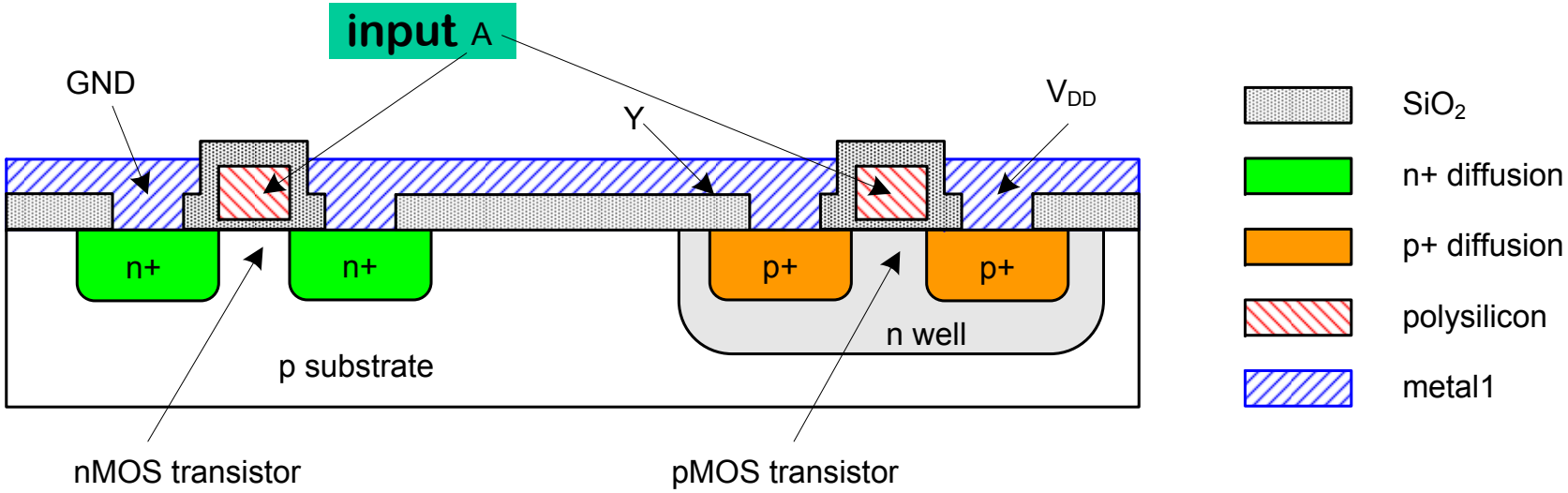


CMOS Fabrication

- CMOS transistors are fabricated on silicon wafer
- Lithography process similar to printing press
- On each step, different materials are deposited or etched
- Easiest to understand by viewing both top and cross-section of wafer in a simplified manufacturing process

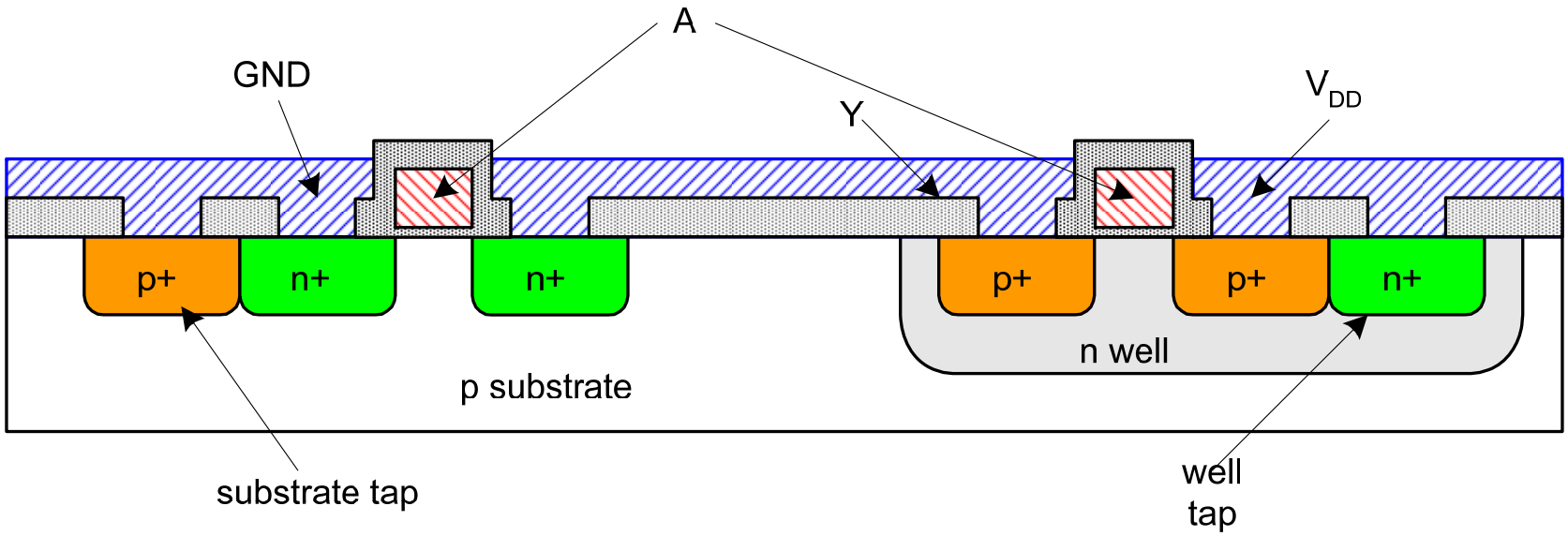
Inverter Cross-section

- **Typically use P-type substrate for NMOS transistors**
- **Requires N-well for body of PMOS transistors**
  - Silicon dioxide (SiO_2) prevents metal from shorting to other layers



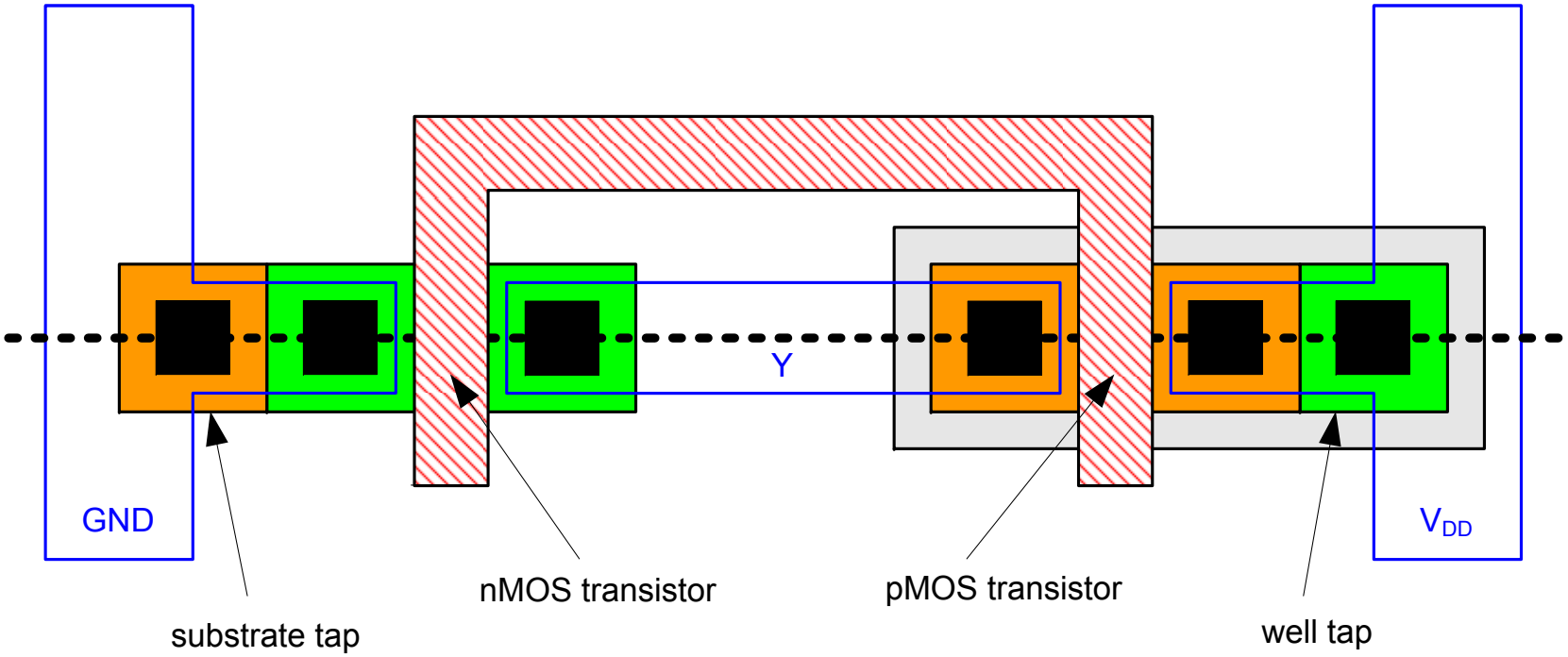
Well and Substrate Taps

- **P-type substrate (body) must be tied to GND**
- **N-well is tied to V_{DD}**
- **Use heavily doped well and substrate contacts (taps)**
  - Establish a good *ohmic contact* providing low resistance for bidirectional current flow



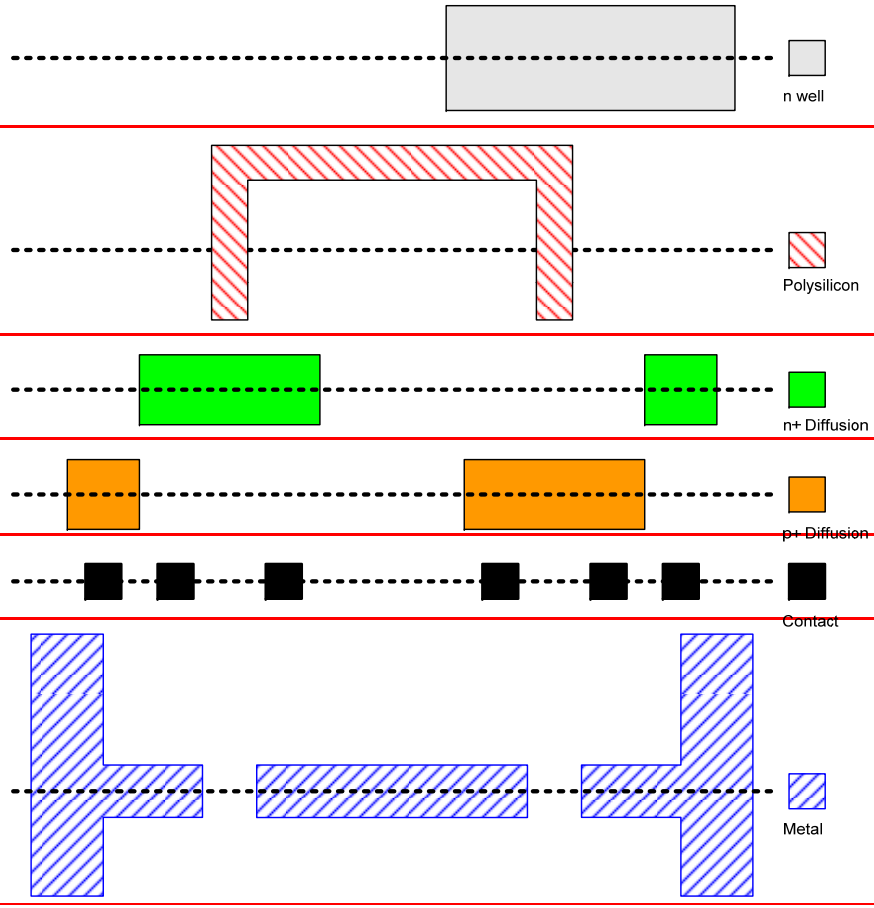
Inverter Mask Set

- **Transistors and wires are defined by *masks***
  - Inverter can be obtained using six masks: n-well, polysilicon, n+ diffusion, p+ diffusion, contacts and metal
- **Cross-section taken along dashed line**



Detailed Mask Views

- **Six masks**
  - n-well
  - Polysilicon
  - N+ diffusion
  - P+ diffusion
  - Contact
  - Metal



Fabrication

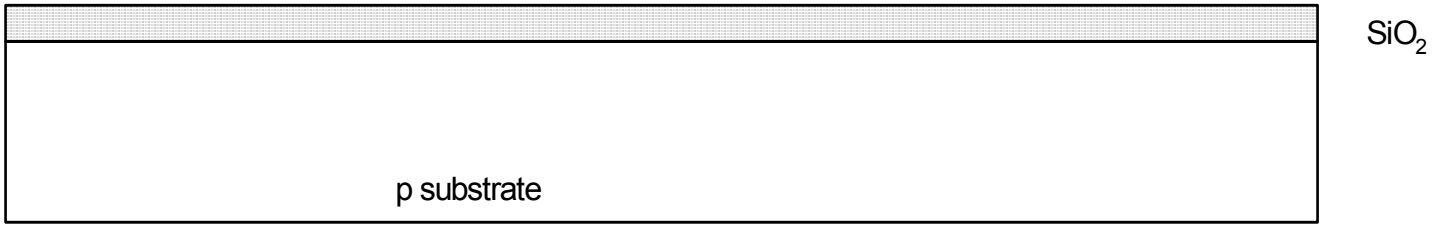
- Chips are built in huge factories called fabs
- Contain clean rooms as large as football fields

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

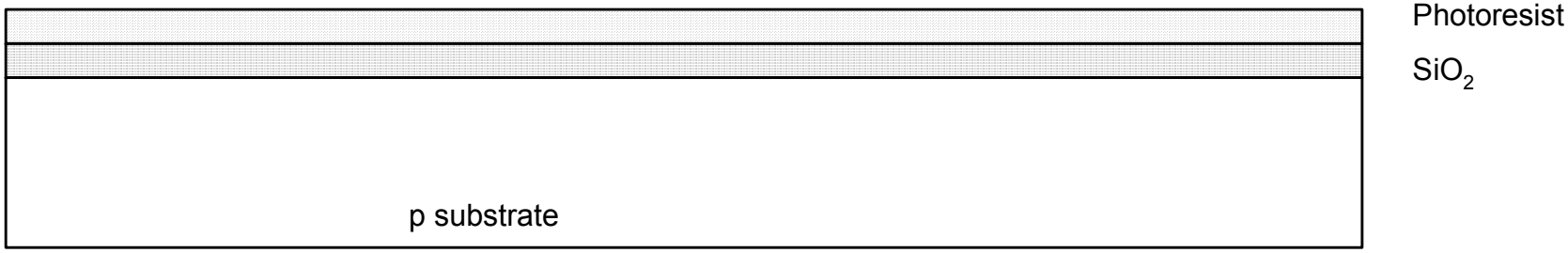
Oxidation

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



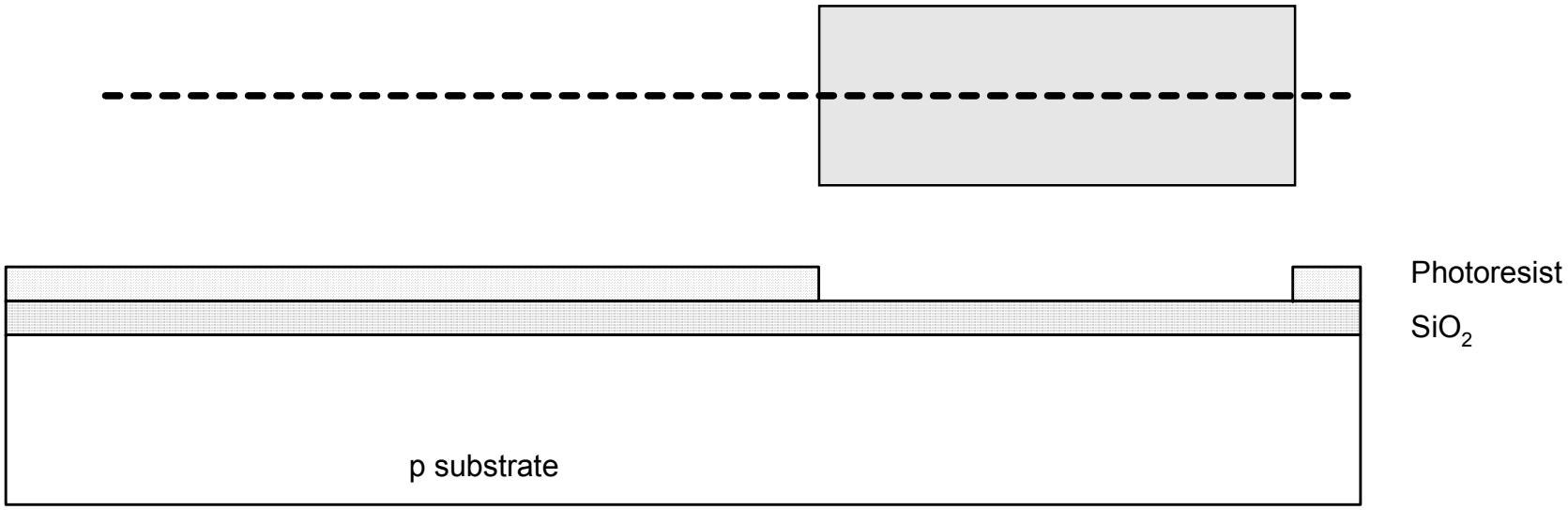
Photoresist

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



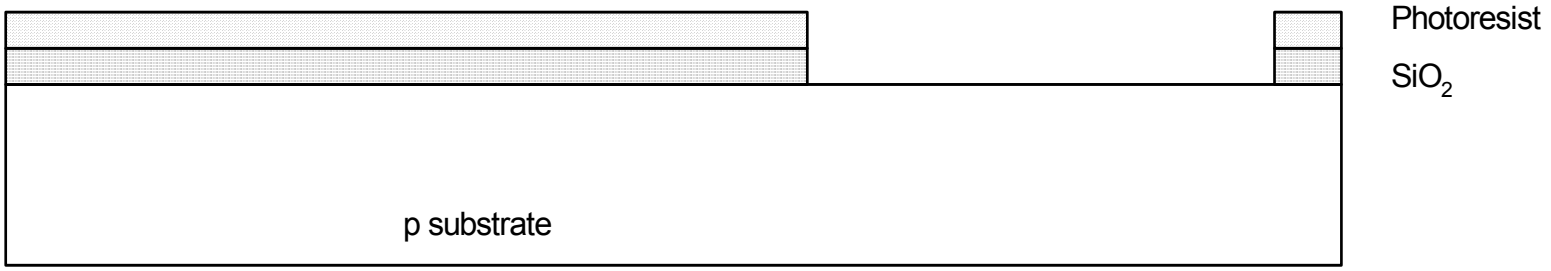
Lithography

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



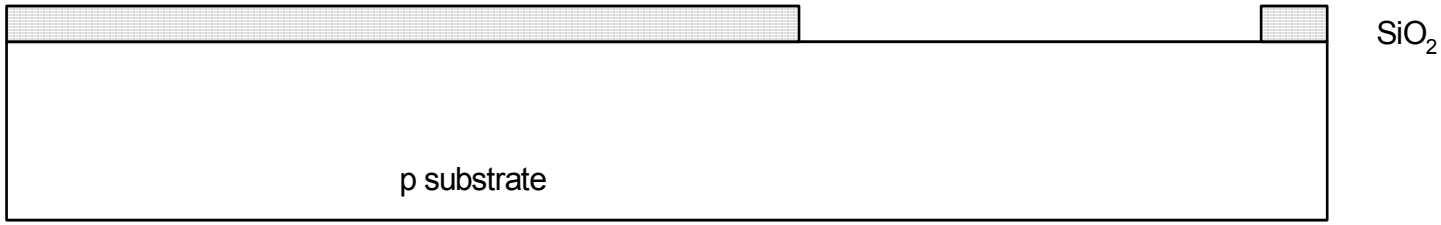
Etch

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



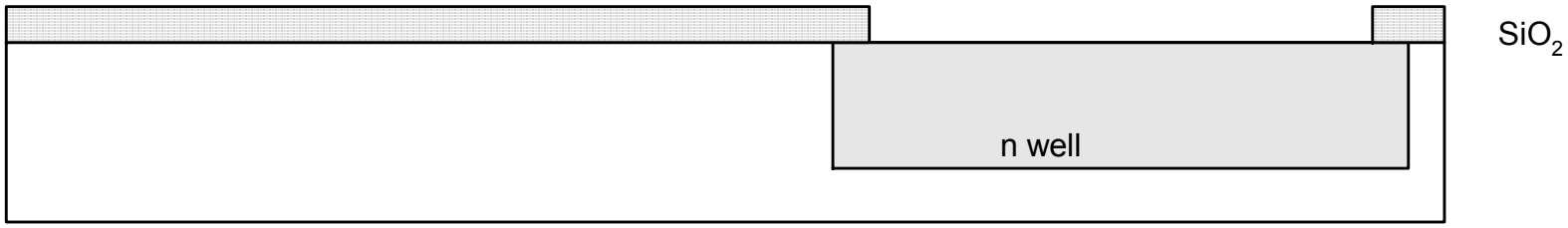
Strip Photoresist

- **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 SiO_2, only enter exposed Si

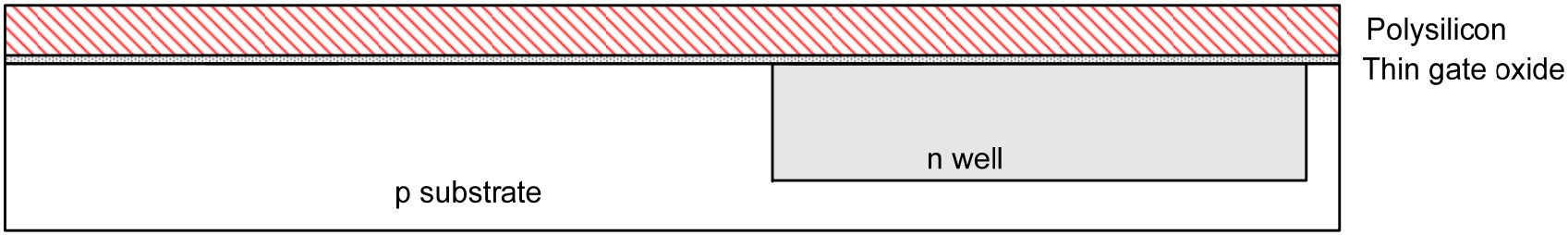


Strip Oxide

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

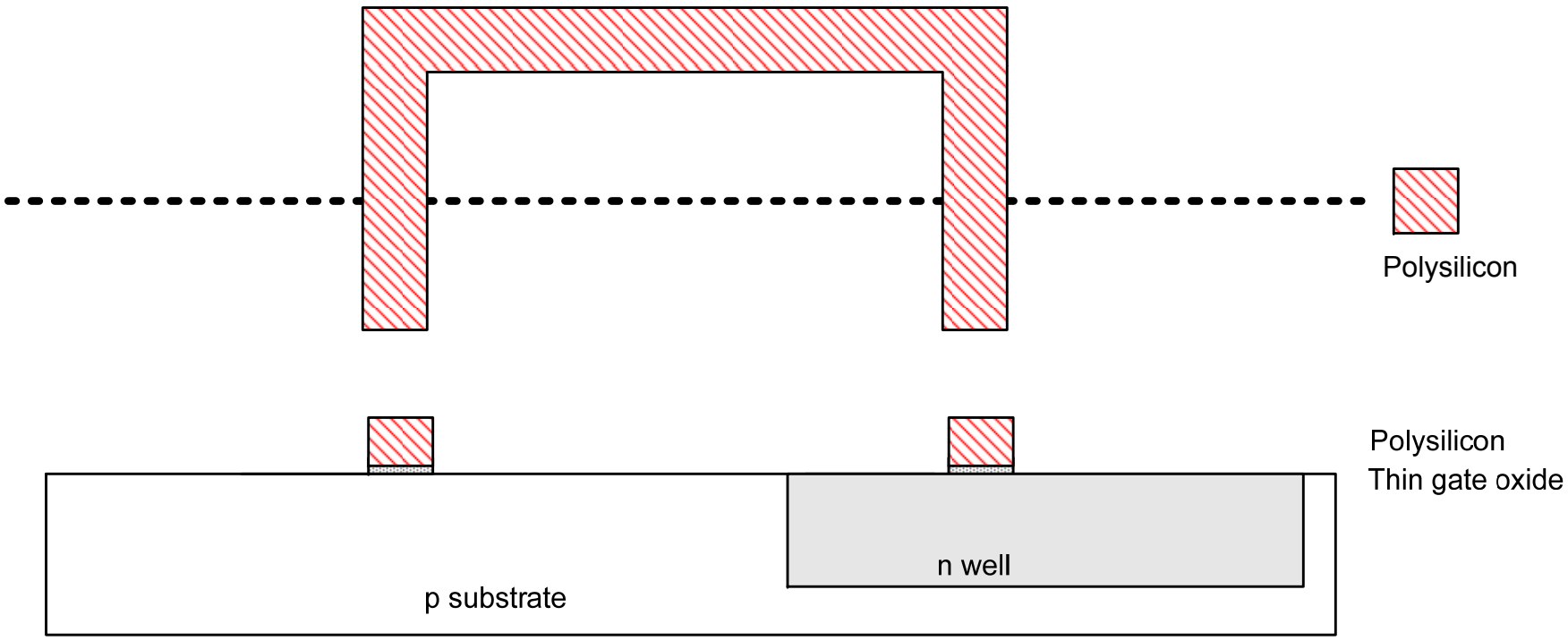
Polysilicon

- **Deposit very thin layer of gate oxide (SiO_2)**
  - < 20 Å (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



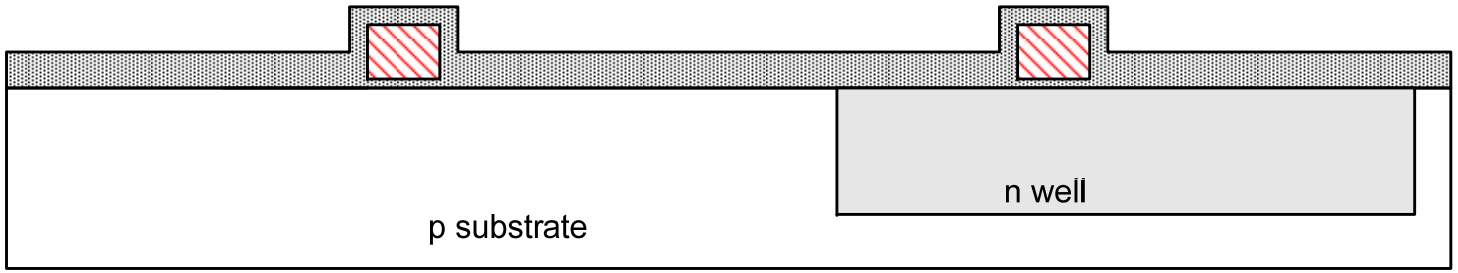
Polysilicon Patterning

- **Use same lithography process to pattern polysilicon**



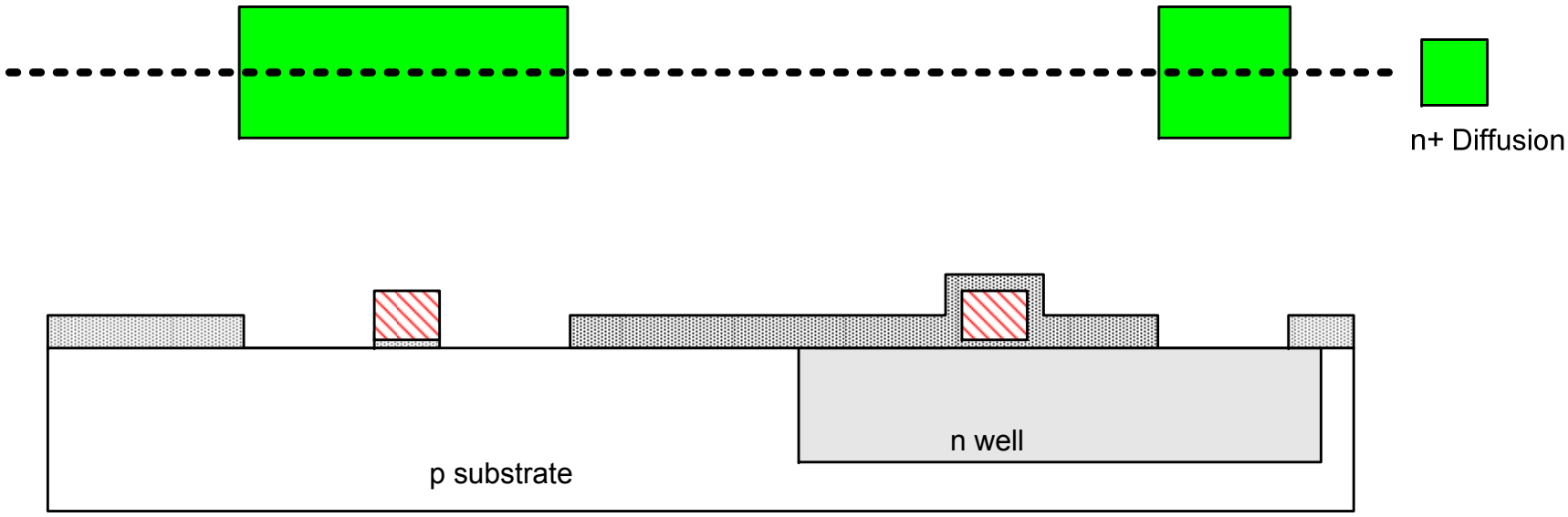
Self-Aligned Process

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



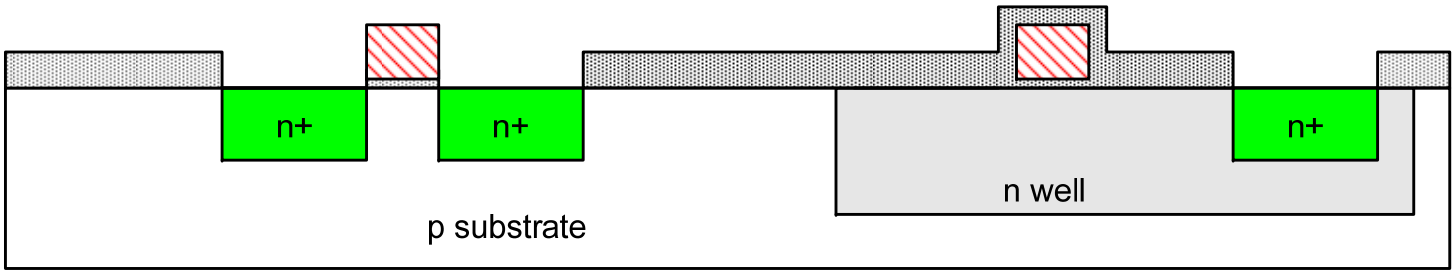
N-diffusion

- 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



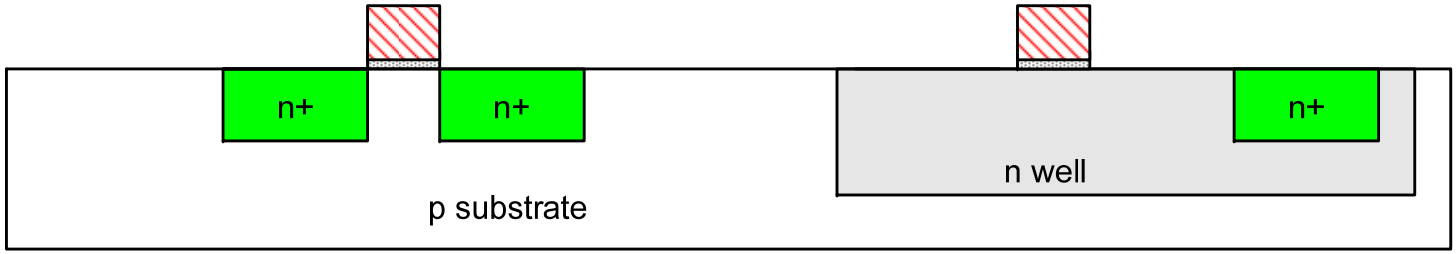
N-diffusion cont.

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



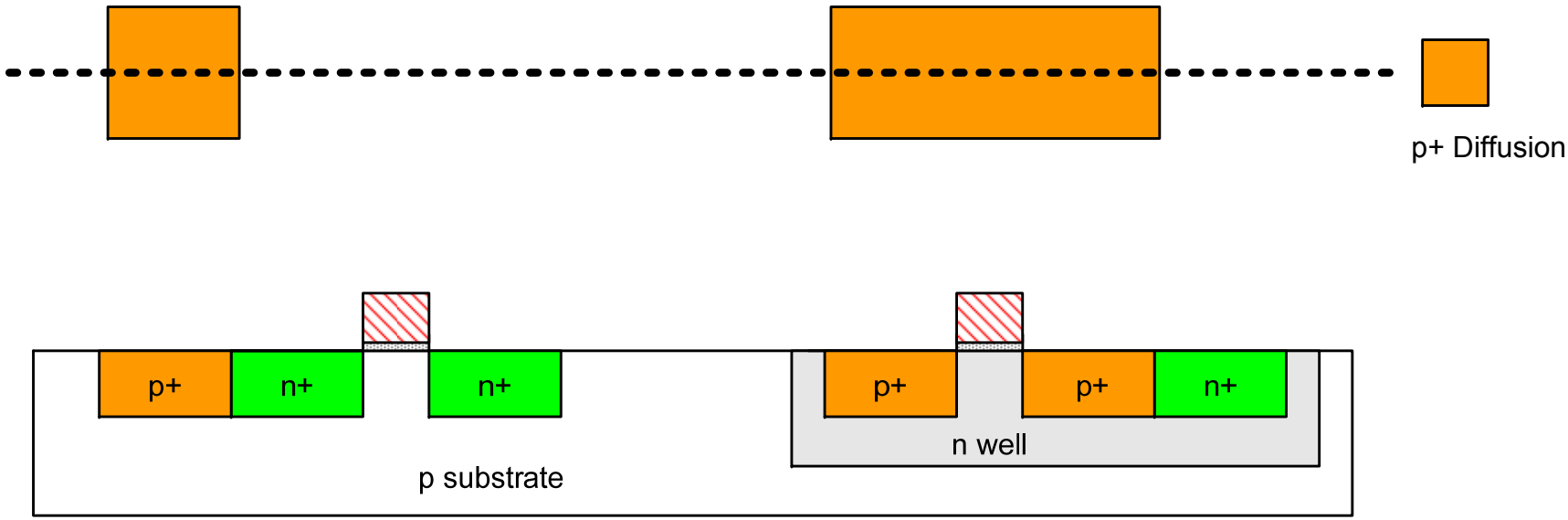
N-diffusion cont.

- **Strip off oxide to complete patterning step**



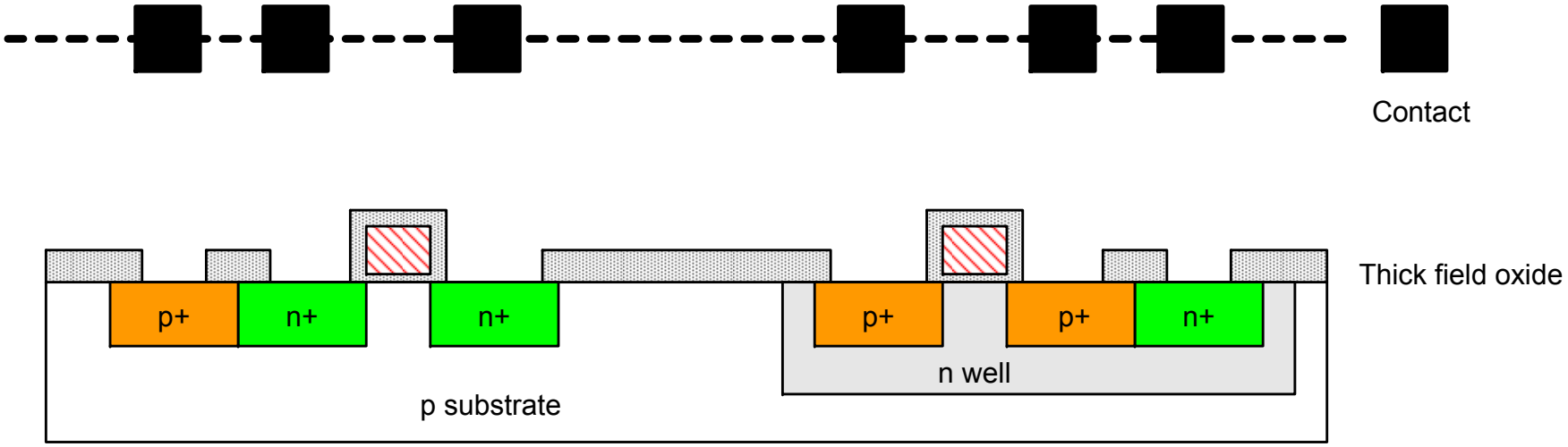
P-Diffusion

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



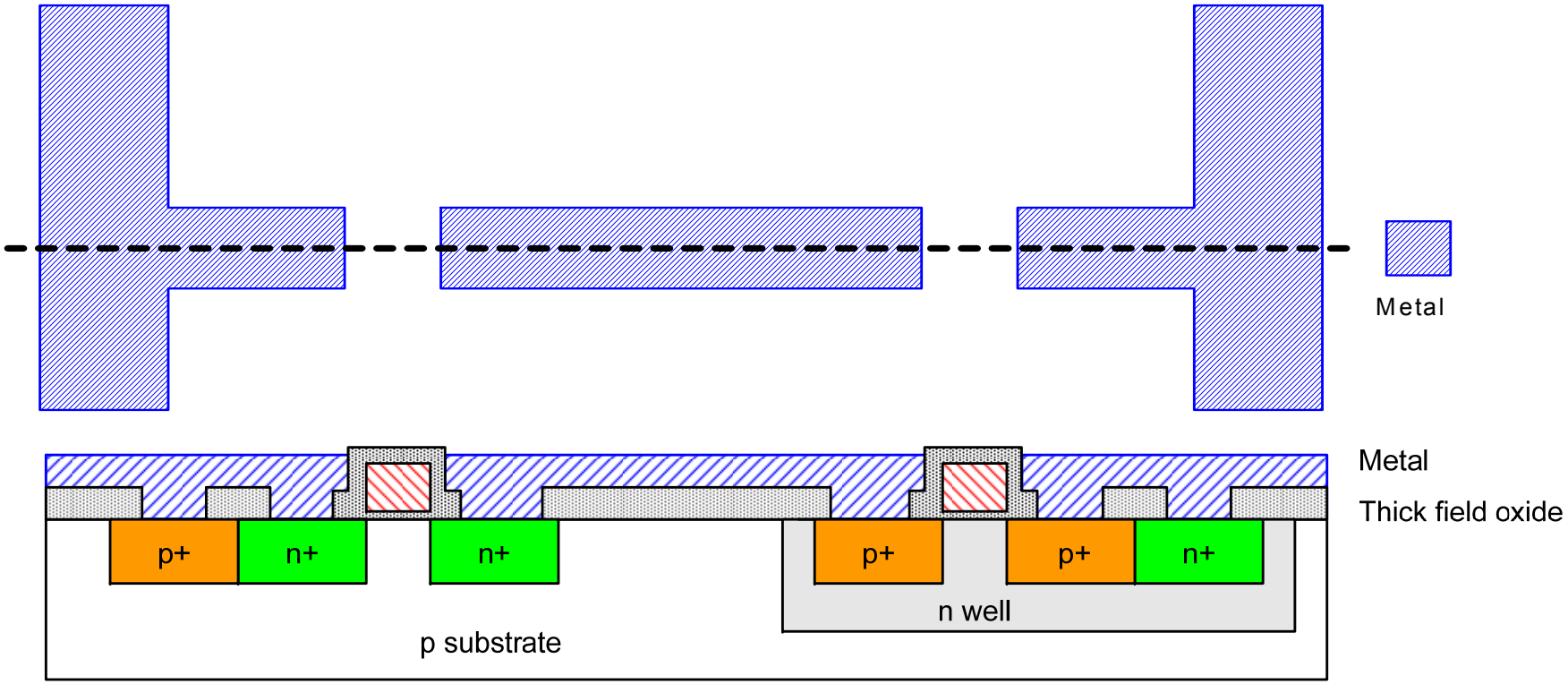
Contacts

- 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

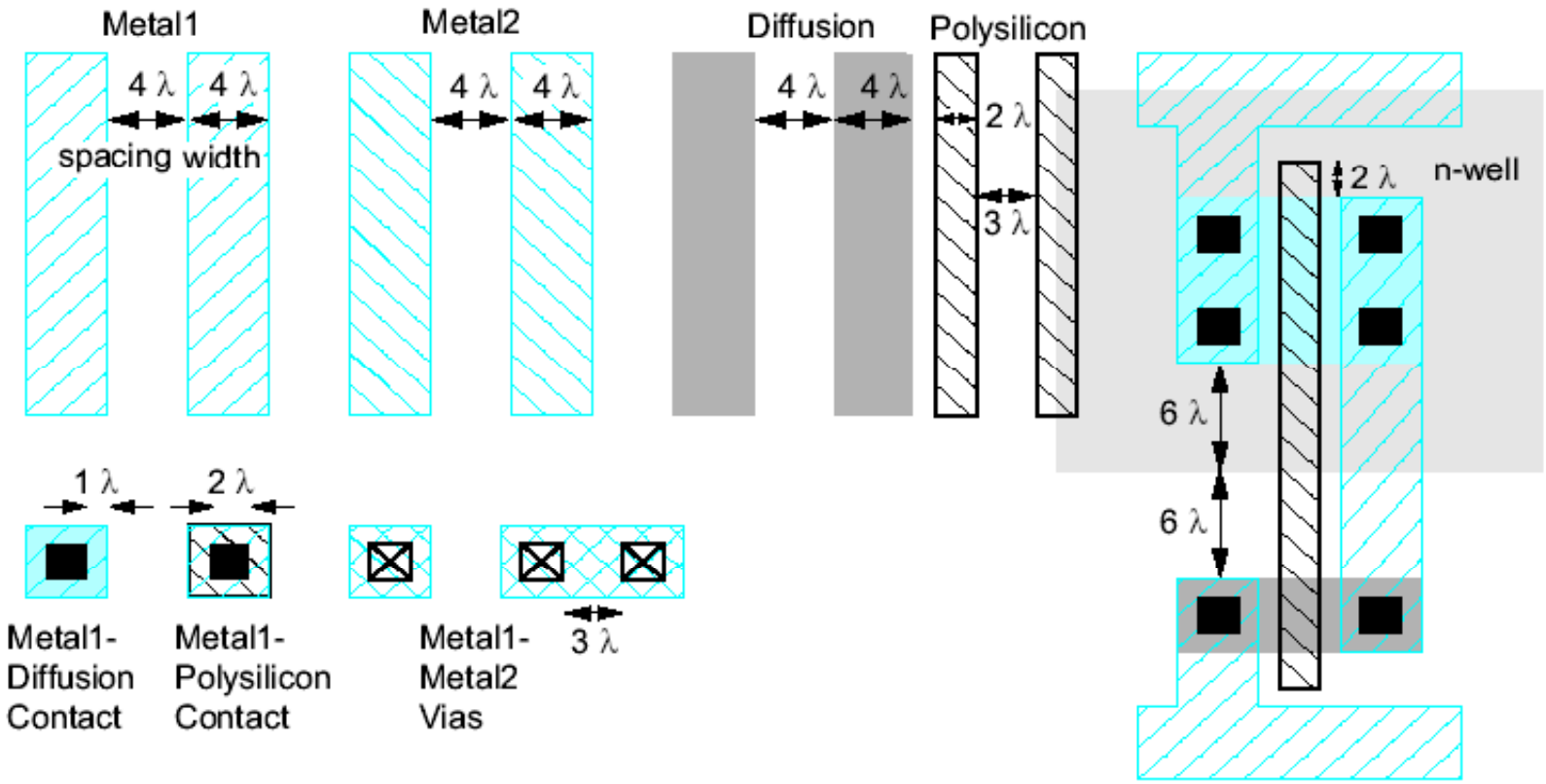


Layout

- Chips are specified with set of masks
- 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 improves 30% every 3 years or so
- Normalize for feature size when describing design rules
- Express rules in terms of $\lambda = f/2$
  - E.g. λ = 0.3 μm in 0.6 μm process

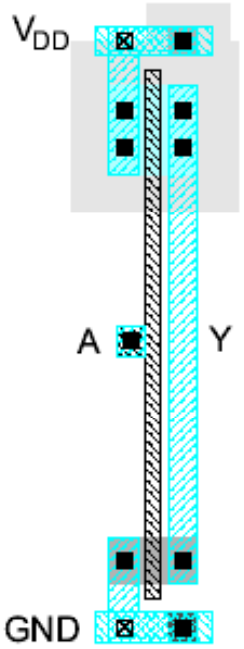
Simplified Design Rules

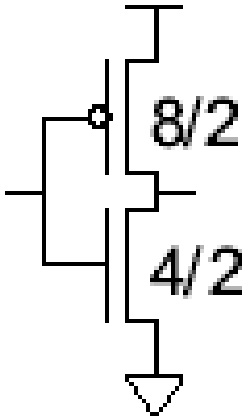
- **Conservative rules to get you started**

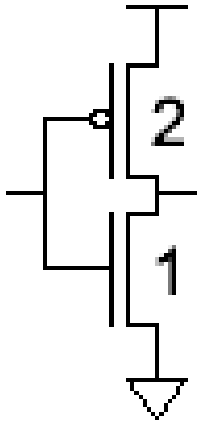


Inverter Layout

- **Transistor dimensions specified as Width / Length**
  - Minimum size is 4λ / 2λ, sometimes called 1 unit
  - In *f* = 0.6 μm process, this is 1.2 μm wide, 0.6 μm long







Summary

- MOS Transistors are stack of gate, oxide, silicon
- Can be viewed as electrically controlled switches
- Build logic gates out of switches
- Draw masks to specify layout of transistors
- Now you know everything necessary to start designing schematics and layout for a simple chip!