CMOS Transistor Notation

- A CMOS transistor has 3 terminals, called the *gate*, *source*, and *drain*
- V_{AB} is the voltage between nodes A and B in a circuit
- Positive power supply (power source)
  - In TTL, written V_{CC} (usually written VCC)
  - In NMOS and CMOS, written V_{DD} (also VDD)
- Negative power supply (power sink)
  - In TTL, written GND ("ground")
  - In NMOS and CMOS, sometimes written V_{SS} (also VSS)
- CMOS uses positive logic: VDD is logic "1", VSS is logic "0"

CMOS Transistors as Switches

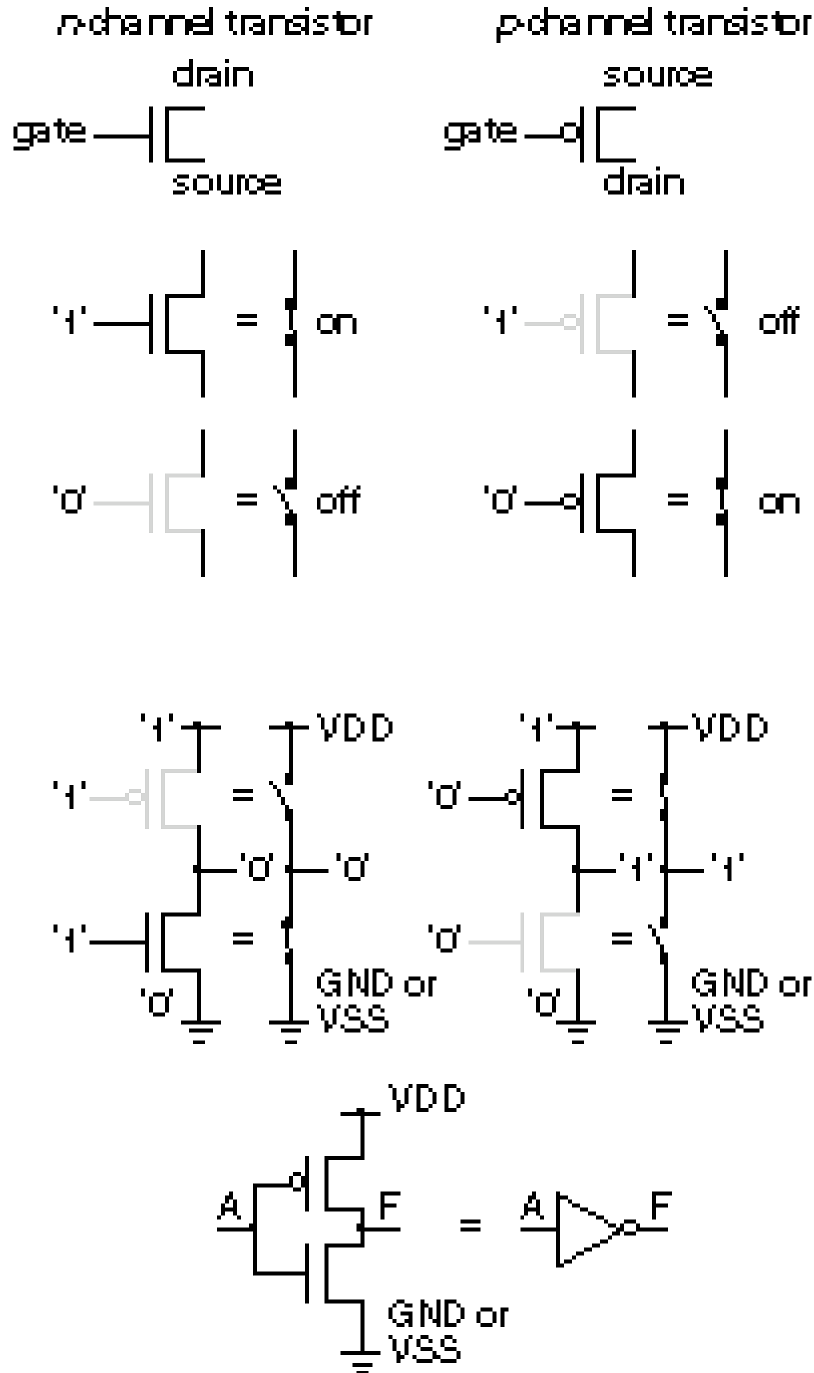

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

CMOS NAND and NOR Gates

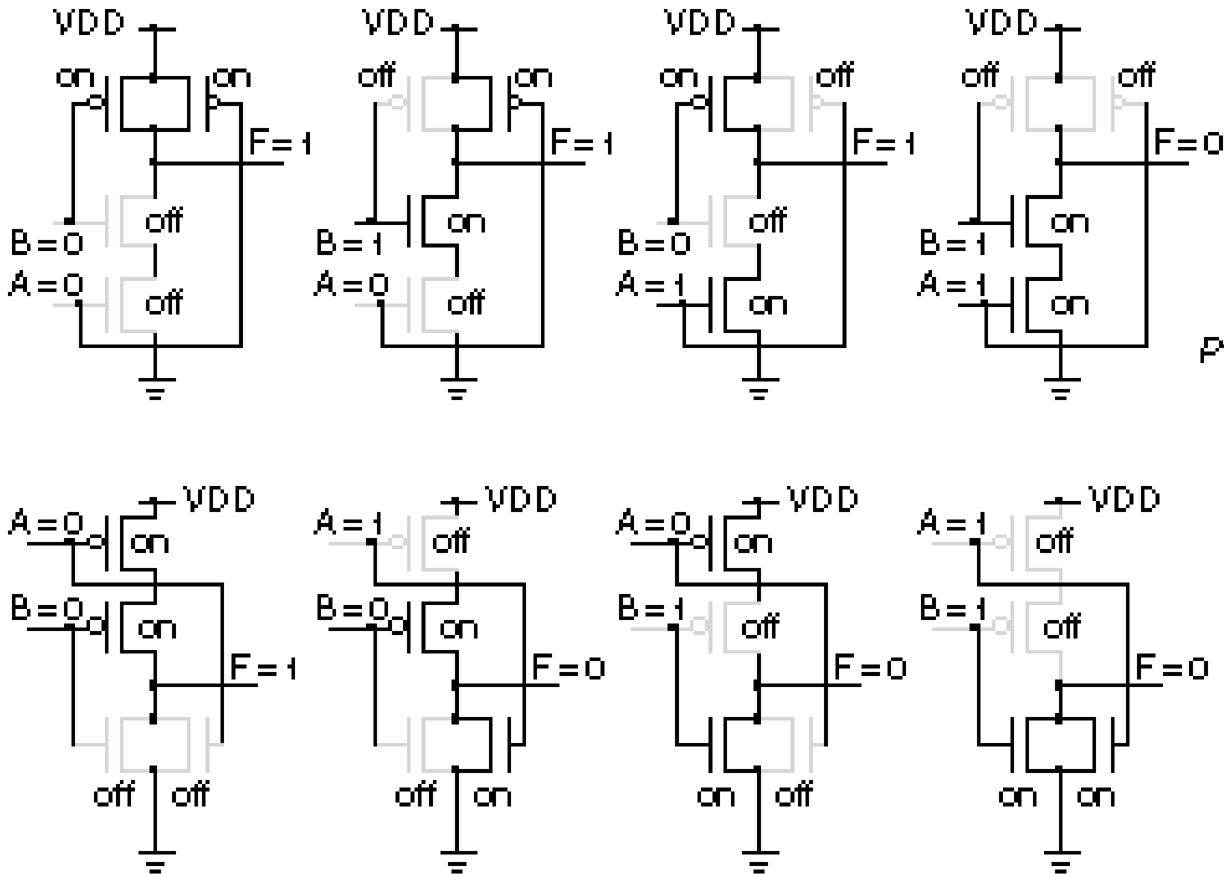

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

- Use two transistors to make a CMOS inverter (as shown on previous slide)
- Use four transistors to make a CMOS 2-input NAND gate
  - Rule of thumb: 1 gate = 4 transistors

IC Fabrication Technologies (Implementing an Inverter)

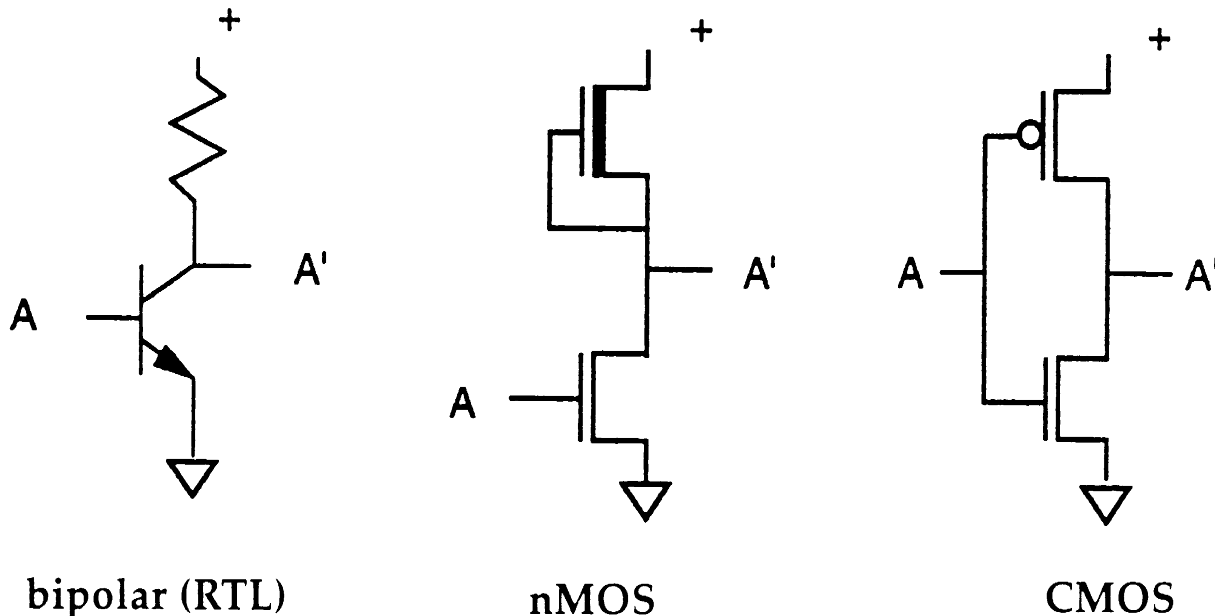

Figure from *Modern VLSI Design*, Wolf, Prentice Hall, 1994

- Bipolar — transistor & resistor (**fastest**)
- NMOS — n-channel depletion mode transistor (top) & n-channel enhancement mode transistor (bottom)
- CMOS — p-channel (**lowest power**) enhancement mode transistor (top) & n-channel enhancement mode transistor (bottom)

Cross-Section of an N-Channel Enhancement Mode MOS Transistor

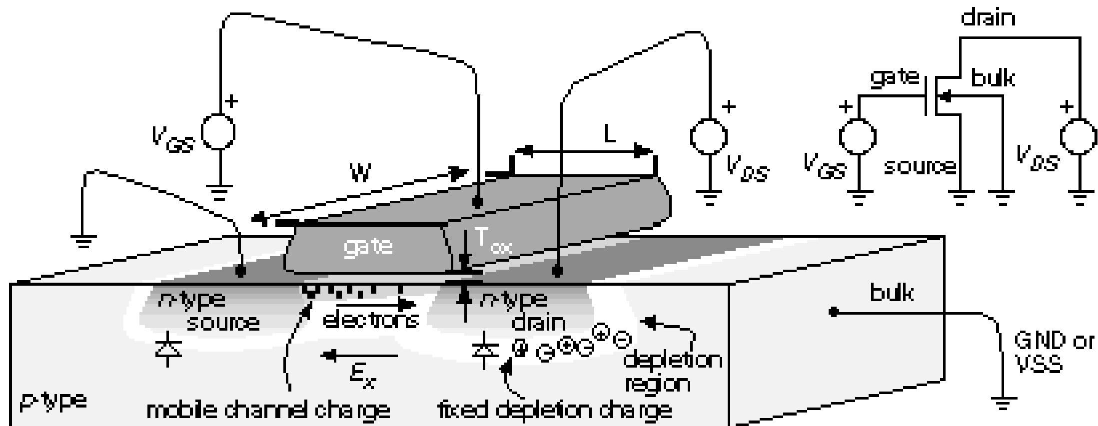

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

- Base is silicon *substrate* (*bulk*, *well*, *tub*) that's been *doped* with p-type impurities (full of positively-charged *holes*)
  - Two *diffusion* areas heavily doped with n-type impurities (full of negatively-charged *electrons*) form the *source* and *drain*
  - Transistor action takes place at the *channel*, connecting the source and drain
- A very thin layer of silicon dioxide (SiO_2), called the *gate oxide*, insulates the *gate*, made of polysilicon, from the channel

Operation of an N-Channel Enhancement Mode MOS Transistor

- Works as a switch — gate-to-source voltage regulates the amount of current that can flow between drain and source
  - When $V_{GS} = 0$, the p-type channel is full of *holes*, and the n-type source and drain contain *electrons*
    - The p-n junctions at source and drain form *diodes* in opposite directions, so no current flows between source and drain
  - As V_{GS} rises above 0, the few n-type impurities that are present in the p-type channel start to attract electrons
    - The electrons migrate toward the (positively charged) gate, but are stopped by the gate oxide, so collect at the top of the channel
  - When V_{GS} rises to the *threshold voltage* (V_t), enough electrons have collected to form an *n-channel inversion layer*, which allows electrons to flow from source to drain (current flows from drain to source)

Operation of an N-Channel Enhancement Mode MOS Transistor (cont.)

- N-channel vs. p-channel
  - N-channel: V_{GS} and V_{DS} both positive, gate and source are n-type (electrons), substrate & channel is p-type (holes), when $V_{GS} >> V_t$ electrons accumulate in channel and flow from source to drain, current flows from drain to source
  - P-channel: V_{GS} and V_{DS} both negative, gate and source are p-type (holes), substrate & channel is n-type (electrons), when V_{GS} >> (more negative) than V_t holes accumulate in channel and flow from source to drain, current flows from source to drain
- Current proportional to W/L of transistor
  - Length (L) = parallel to current flow
  - As W increases, more current can flow
  - As L increases, less current flows

IC Fabrication

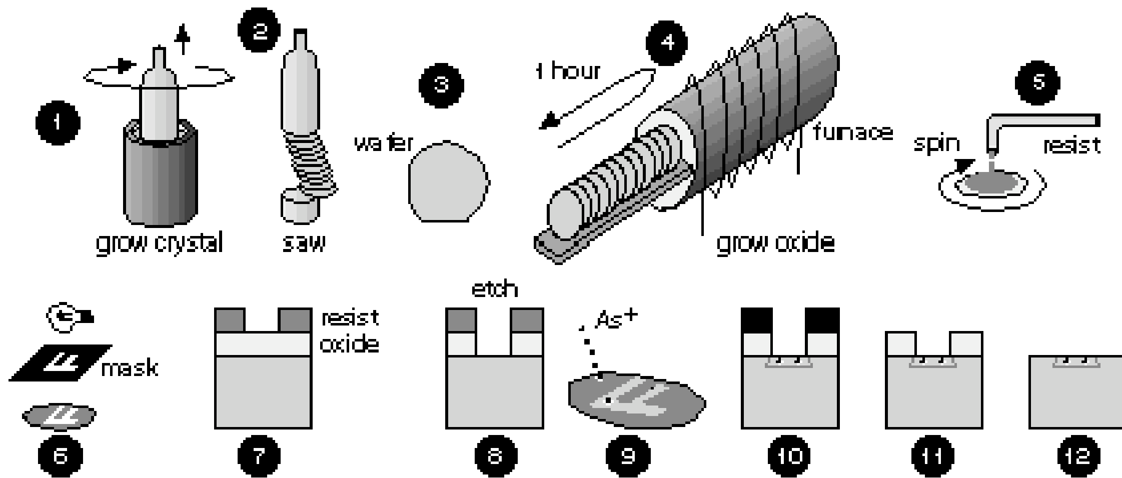

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

1. Start with silicon (Si), refined from quartzite, dope it with p- or n-type impurities, and melt it at 1500°C

2.&3. Draw out a single crystal (6" or 8"), saw it into thin (600μm) *wafers*, polish one side, and grind down an edge or two

4. Batch of wafers (a *wafer lot*) is placed on a *boat* and put in a furnace to grow a layer (1000 Å) of silicon dioxide (SiO_2) (called the *oxide*)

IC Fabrication (cont.)

- IC fabrication process uses a series of masking steps to create the layers that form the transistors etc. on the chip

5. A thin layer (1 μm) of liquid photoresist (*resist*) is spun onto each wafer, and it is baked at 100°C to harden the resist

6. The wafer is partially exposed to ultraviolet light through a *mask*, which polymerizes the exposed areas; the polymerized part is then removed using an organic solvent

7. The exposed oxide is *etched* away, making the oxide match the mask

8. The exposed silicon substrate is doped with appropriate ions by an *ion implanter*

9. & 10. Resist and oxide are removed

Fabricating a CMOS Transistor

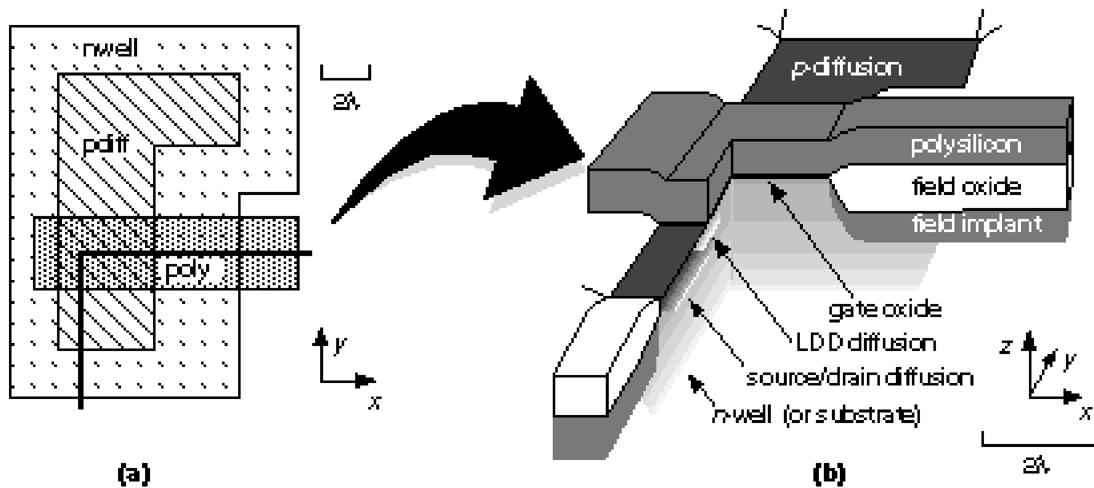

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

- Continuing the fabrication process:
  - Polycrystalline silicon (*poly*) is deposited using using *chemical vapor deposition* (*CVD*) to deposit dopants using a gaseous source in a furnace
    - Poly wires (e.g., transistor gates) are deposited before diffusion to create self-aligned transistors — this avoids small gaps that might otherwise occur if the order is reversed
  - Metal layers are deposited in a similar manner, called *sputtering*

Fabricating a CMOS Transistor (cont.)

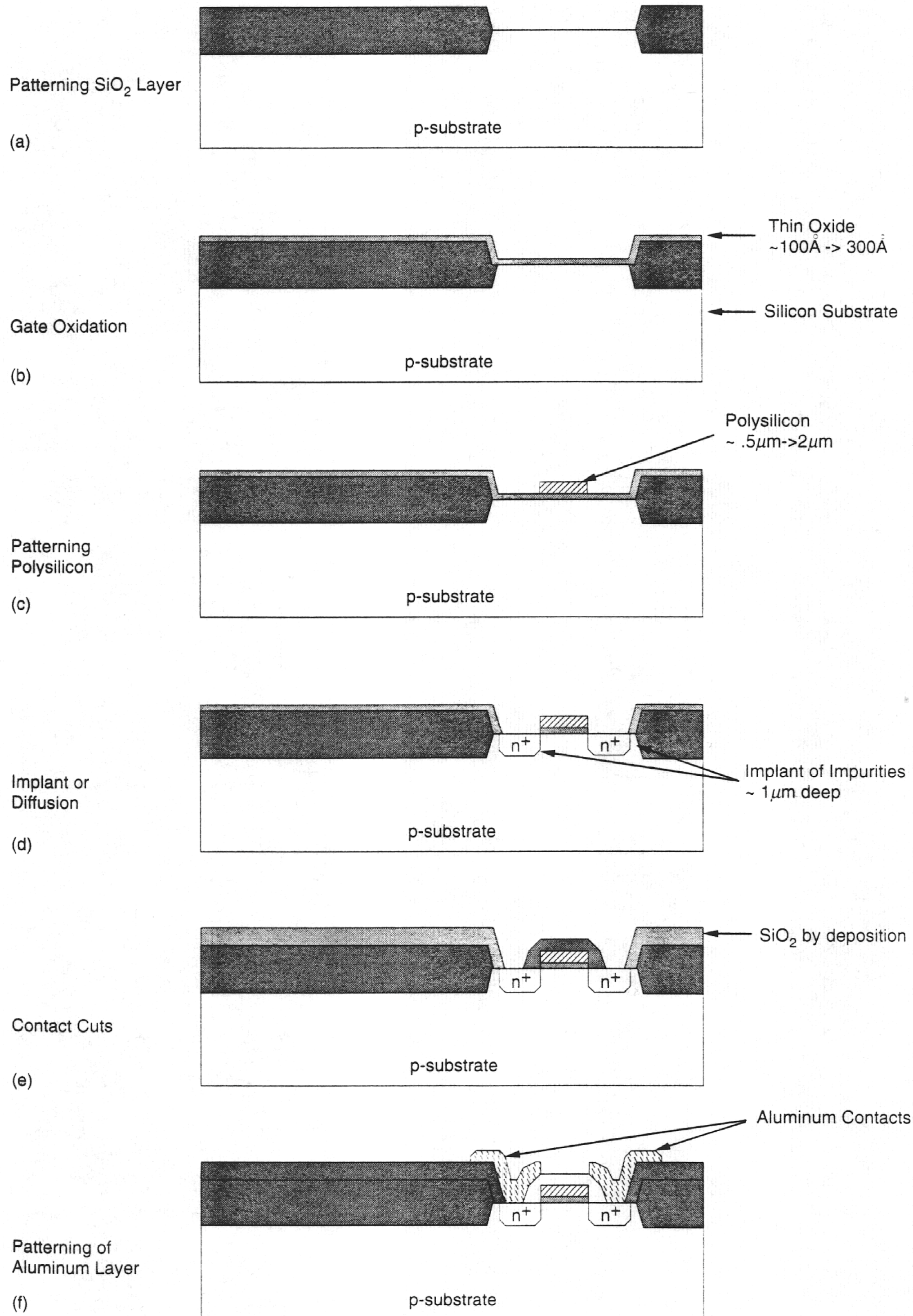

Figure from *Principles of CMOS VLSI Design*, Weste, Addison-Wesley, 1993

Wires and Vias

- "Wires" can be fabricated using diffusion, polysilicon, or metal
  - Must be insulated from each other using silicon dioxide; can be built up as layers
  - Diffusion used only for within a cell
  - Poly used between adjacent cells
  - Metal used for longer connections
    - Chip may have 5-6 layers of metal
  - Cuts in the silicon dioxide between layers are called *vias*

- Metal layer 1 is used for VDD and VSS; other layers are used for interconnection
  - Under high currents, electron collisions with metal grains can cause the metal to move (*metal migration*), so in large designs, sizing the power supply lines is critical to keep chip from failing

CMOS Inverter in a n-well Process

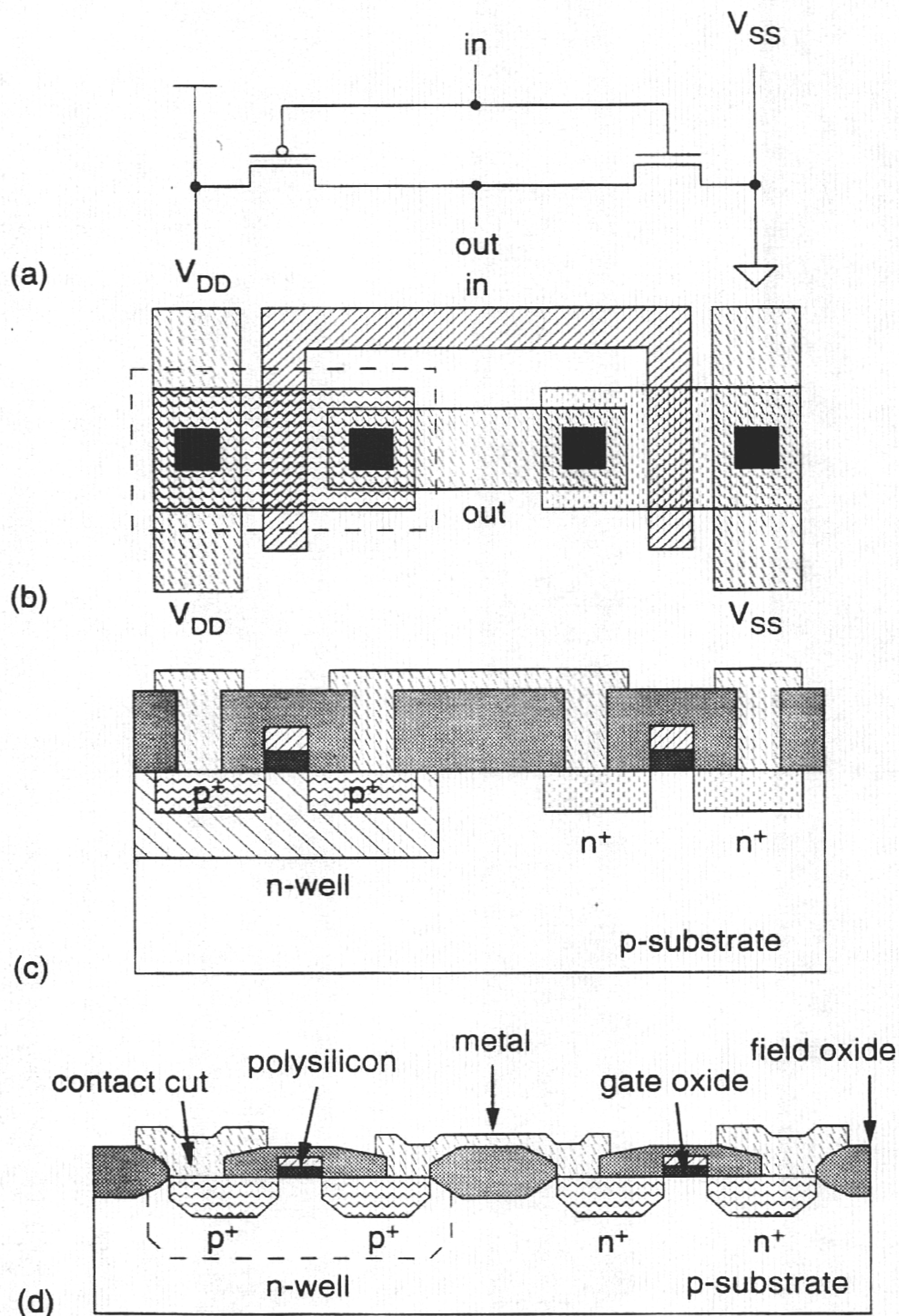

Figure from *Principles of CMOS VLSI Design*, Weste, Addison-Wesley, 1993

CMOS Inverter in a n-well Process, With Tub Ties

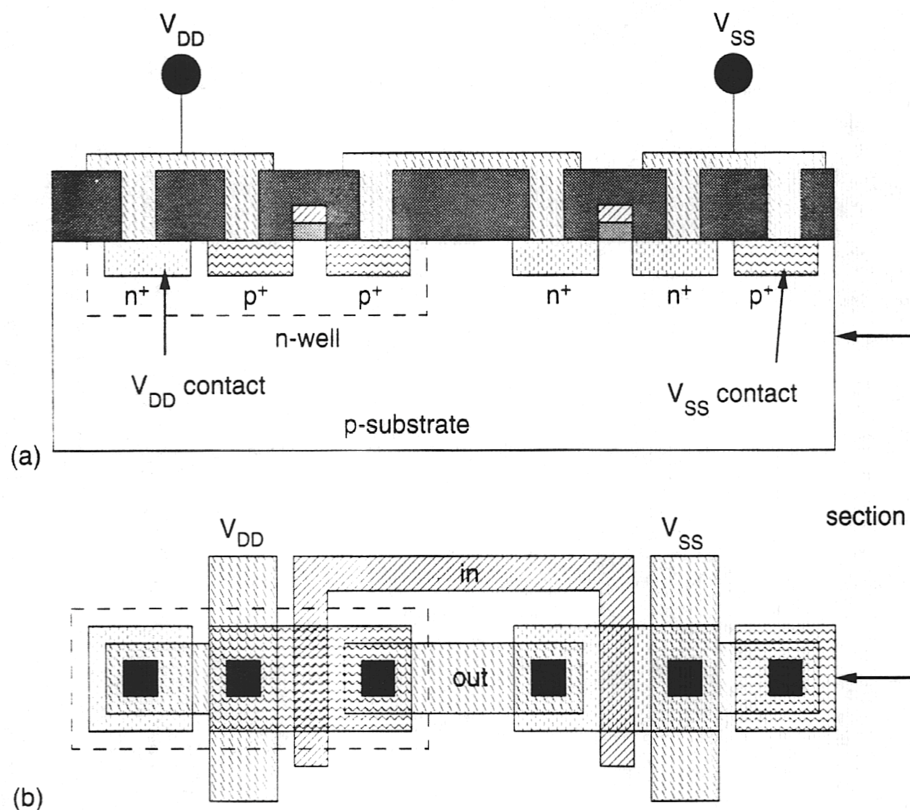

Figure from *Principles of CMOS VLSI Design*, Weste, Addison-Wesley, 1993

- An n-channel (p-channel) transistor's substrate must be tied to VSS (VDD)
  - The special vias that provide these connections are called *tub ties*
    - Need one every 1-2 transistors (SCMOS)
  - Tie area in substrate is heavily doped to provide a low-resistance connection

Design Rules & Fabrication Errors

- Common fabrication errors:
  - Wire too wide — may *short* (contact) an adjacent wire
  - Wire too narrow — may break under load, and become *open*

- Solution — impose *design rules* to specify what's legal and illegal
  - Wires — specify minimum width and minimum spacing between wires
  - Poly — must extend beyond channel to ensure that there is no short between source and drain
  - Diffusion — must extend enough to have room for a contact to that region
  - Via — must be smaller than what it's contacting, what it's contacting must extend back under SiO_2

Scaleable Design Rules (1997)

- Fabrication processes are constantly being improved
  - (Gordon) Moore's Law (version 2) says that the number of transistors on a chip is doubling every 18 months
- We take advantage of these improvements by designing according to *scaleable design rules*
  - Specified in terms of λ, the minimum feature size possible in that process
  - In MOSIS SCMOS rules, mimium channel width (poly) is 2λ , and minimum wire width is 2λ
    - MOSIS = MOS Implementation Service, located at the Information Sciences Institute at the University of Southern California (USC), does small-volume fabrication for universities (partially NSF-funded) and commercially (www.isi.edu)

SCMOS Design Rules (1997) (cont.)

- MOSIS (rev. 7), dimensions in λ

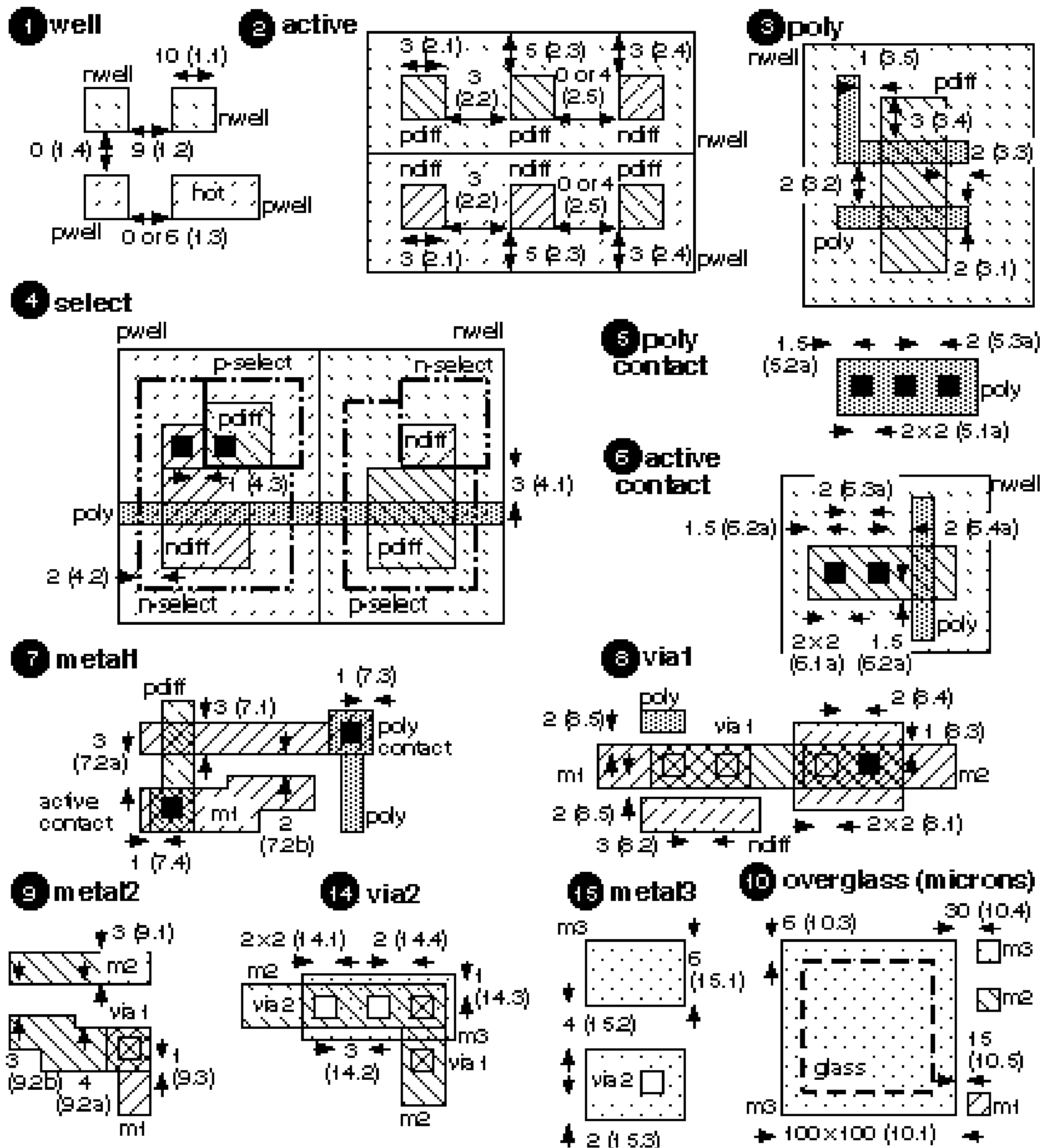

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

SCMOS Design Rules (1997) (cont.)

■ MOSIS (rev. 7), dimensions in λ

TABLE 2.7 MOSIS scalable CMOS rules version 7the process front end.

Layer	Rule	Explanation	Value / λ
well (CWN, CWP)	1.1	minimum width	10
	1.2	minimum space (different potential, a hot well)	9
	1.3	minimum space (same potential)	0 or 6
	1.4	minimum space (different well type)	0
active (CAA)	2.1/2.2	minimum width/space	3
	2.3	source/drain active to well edge space	5
	2.4	substrate/well contact active to well edge space	3
	2.5	minimum space between active (different implant type)	0 or 4
poly (CPG)	3.1/3.2	minimum width/space	2
	3.3	minimum gate extension of active	2
	3.4	minimum active extension of poly	3
	3.5	minimum field poly to active space	1
select (CSN, CSP)	4.1	minimum select spacing to channel of transistor[1]	3
	4.2	minimum select overlap of active	2
	4.3	minimum select overlap of contact	1
	4.4	minimum select width and spacing[2]	2
poly contact (CCP)	5.1.a	exact contact size	2 ∞ 2
	5.2.a	minimum poly overlap	1.5
	5.3.a	minimum contact spacing	2
active contact (CCA)	6.1.a	exact contact size	2 ∞ 2
	6.2.a	minimum active overlap	1.5
	6.3.a	minimum contact spacing	2
	6.4.a	minimum space to gate of transistor	2

Table from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

Economics of ASICs

- For a given design, which type of ASIC is the most cost-effective?
  - (full-custom) ASIC?
  - MGA (mask-programmable gate array)?
  - CBIC (cell-based integrated circuit = standard-cell-based ASIC)?

- Answer: consider the ASIC as a product, and examine the fixed costs and variable costs
  - total product cost =
    fixed product cost + variable product cost
  - *Fixed product cost* is independent of sales volume
    - Fixed product costs amortized per product sold decrease as sales volume increases
  - *Variable product cost* includes assembly costs and manufacturing costs

Example of ASIC Economics (Warning — 1997 Numbers!)

- Sample costs:
  - CBIC: fixed cost $146,000; part cost $8
  - MGA: fixed cost $86,000; part cost $10
  - FPGA: fixed cost $21,800; part cost $39

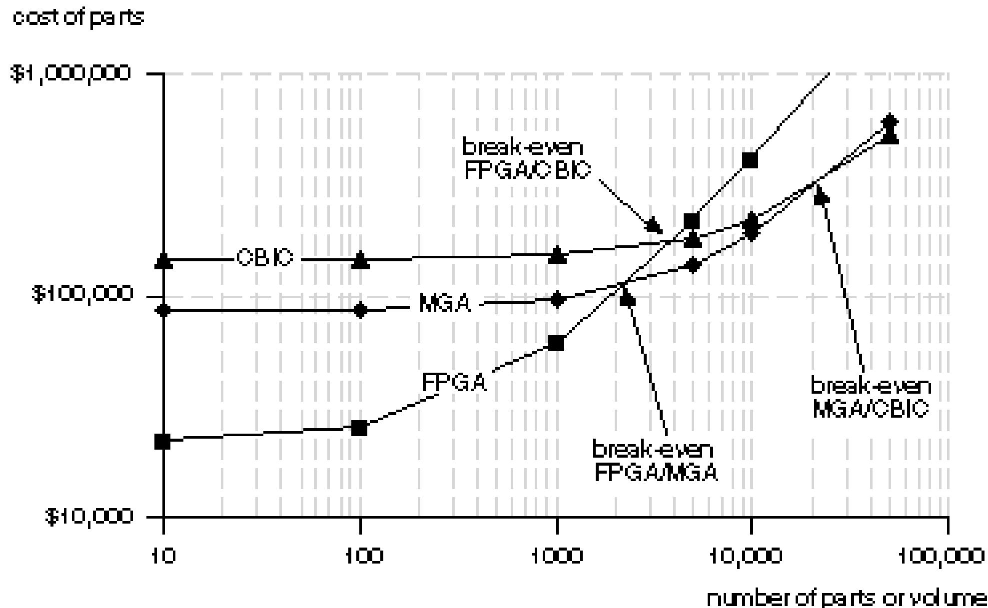

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

- Break-even points:
  - FPGA to MGA is around 2,000 parts
  - FPGA to CBIC is around 4,000 parts
  - MGA to CBIC is around 20,000 parts

ASIC Fixed Costs (1997 Numbers!)

- Design: estimate of designer productivity
- Production test: make sure the IC works
- Non-recurring engineering (NRE): work done by ASIC vendor — developing mask, production tests, prototypes, etc.

	FPGA		MGA		CBIC	
Training:	$800		$2,000		$2,000	
Days		2		5		5
Cost/day		$400		$400		$400
Hardware	$10,000		$10,000		$10,000	
Software	$1,000		$20,000		$40,000	
Design:	$8,000		$20,000		$20,000	
Size (gates)		10,000		10,000		10,000
Gates/day		500		200		200
Days		20		50		50
Cost/day		$400		$400		$400
Design for test:			$2,000		$2,000	
Days				5		5
Cost/day				$400		$400
NRE:			$30,000		$70,000	
Masks				$10,000		$50,000
Simulation				$10,000		$10,000
Test program				$10,000		$10,000
Second source:	$2,000		$2,000		$2,000	
Days		5		5		5
Cost/day		$400		$400		$400
Total fixed costs	$21,800		$86,000		$146,000	

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997

ASIC Variable Costs (1997 Numbers!)

- Wafer size: 6” & 8” common, 12” soon
- 10k gates = small design, 100k = large
- Gate utilization: space used for gates, not used for interconnect
- Defect density is measure of fabrication quality (defect on a die is usually fatal)
- Yield is percentage of usable dies

	FPGA	MGA	CBIC	Units
Wafer size	6	6	6	inches
Wafer cost	1,400	1,300	1,500	$
Design	10,000	10,000	10,000	gates
Density	10,000	20,000	25,000	gates/sq.cm
Utilization	60	85	100	%
Die size	1.67	0.59	0.40	sq.cm
Die/wafer	88	248	365	
Defect density	1.10	0.90	1.00	defects/sq.cm
Yield	65	72	80	%
Die cost	25	7	5	$
Profit margin	60	45	50	%
Price/gate	0.39	0.10	0.08	cents
Part cost	$39	$10	$8	

Figure from *Application-Specific Integrated Circuits*, Smith, Addison-Wesley, 1997