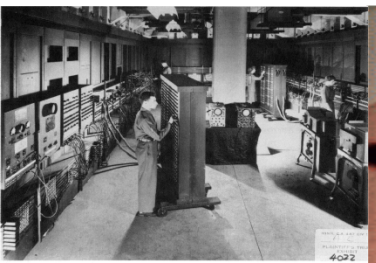
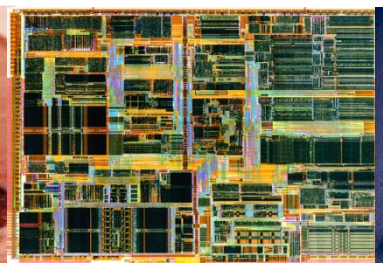
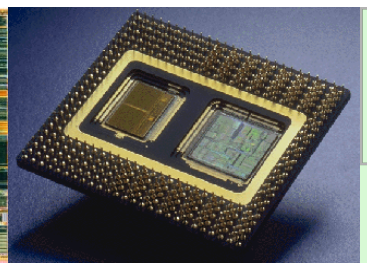
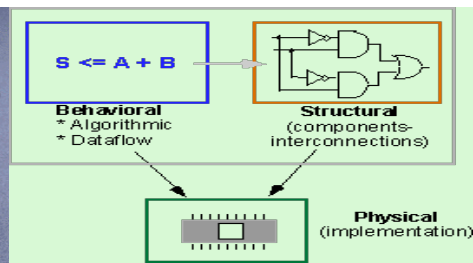
CPE 690: Introduction to VLSI Design

Lecture 5
MOS Transistors and CMOS Logic

Bryan Ackland
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, NJ 07030

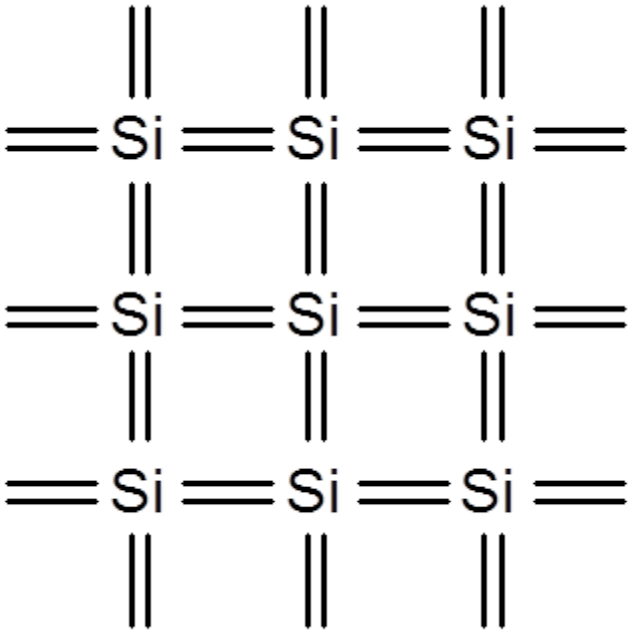
Adapted from Lecture Notes, David Mahoney Harris CMOS VLSI Design

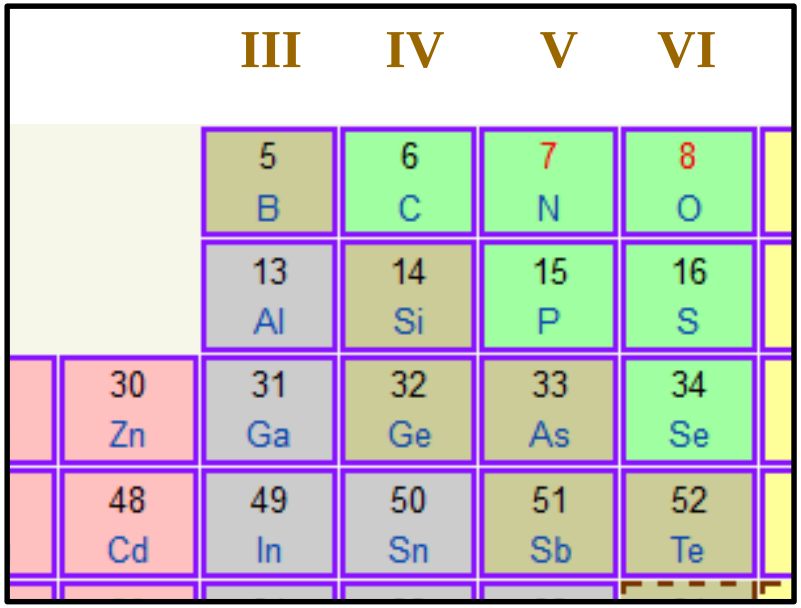


Silicon Lattice

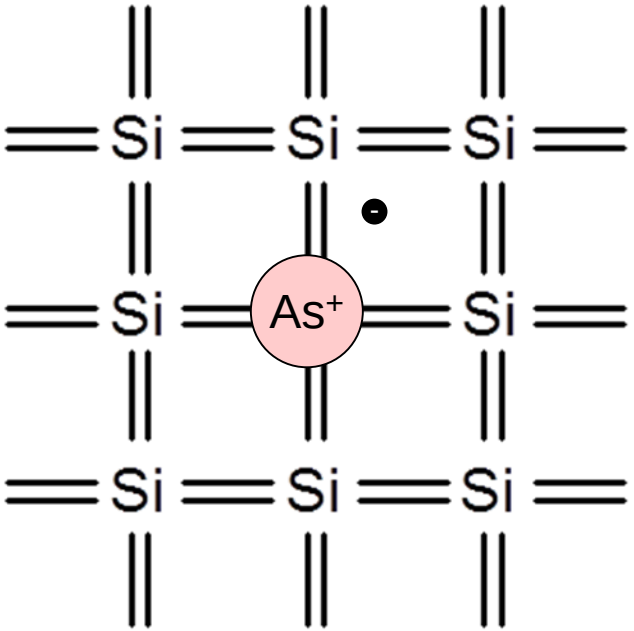
- Transistors are built on a silicon substrate
- Silicon is a Group IV material
- Forms crystal lattice with bonds to four neighbors
- Pure silicon is a semiconductor
- Very few free electrons - conducts very poorly

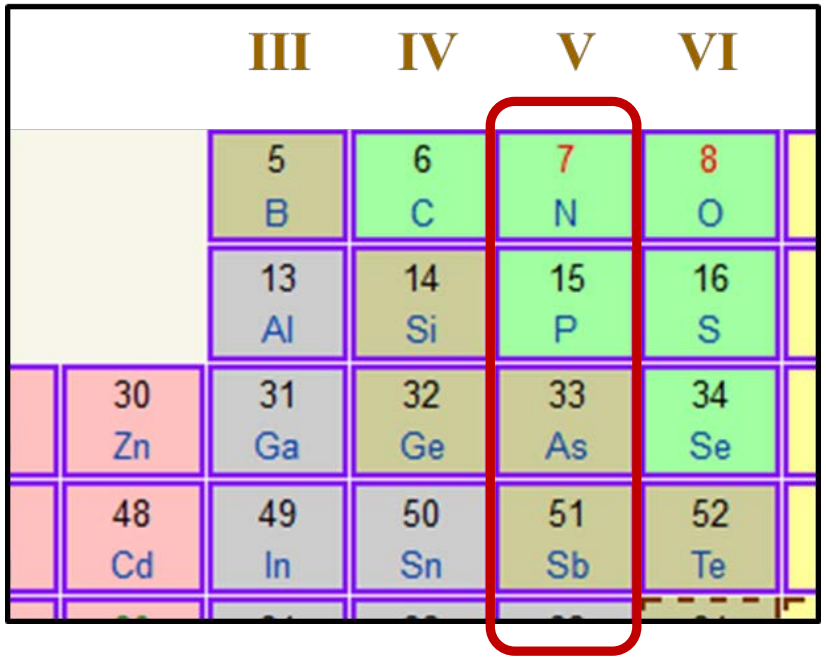




Dopants: Donor Atoms

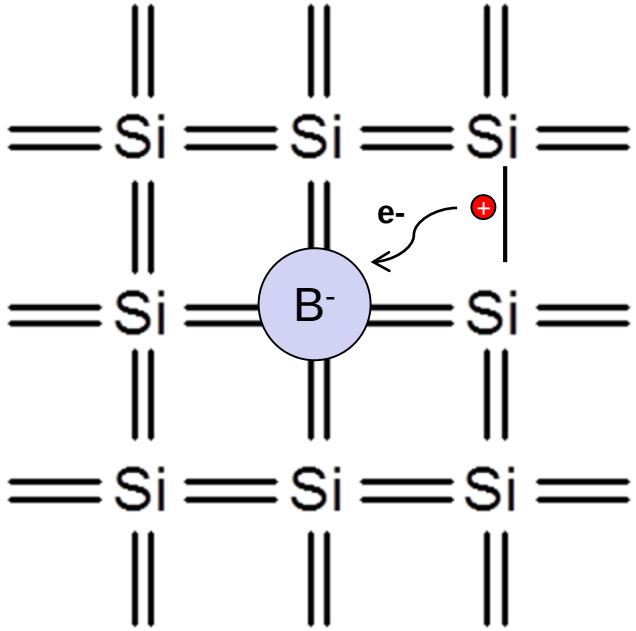
- Adding dopants increases the conductivity
- Group V: extra electron (e.g., Arsenic)
- Extra electron weakly held – free to wander at room temperature
- Semiconductor doped with donors is called n-type

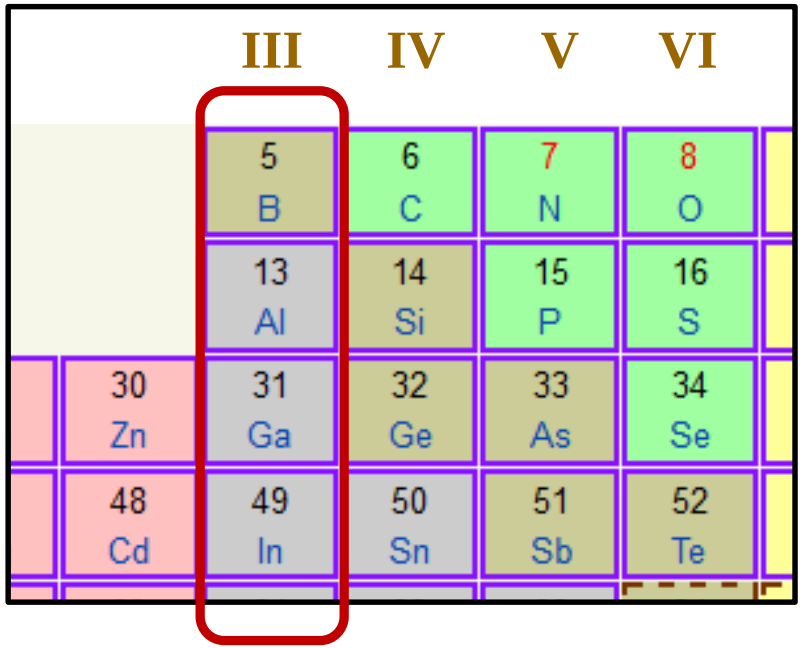




Dopants: Acceptor Atoms

- Group III: missing electron, (known as hole) (e.g, Boron)
- Electron jumps from neighboring silicon atom to fill vacancy – creates a new vacancy (hole)
- Holes can move about as if they were positive carriers (p-type)

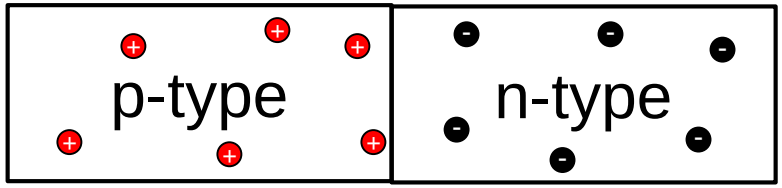




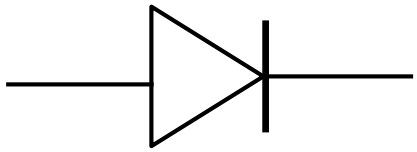
  - if one in million Si atoms is replaced by an acceptor, number of holes available to conduct current increases by a factor of 5×10^6 (same is true for donors and electrons)

PN Junctions

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

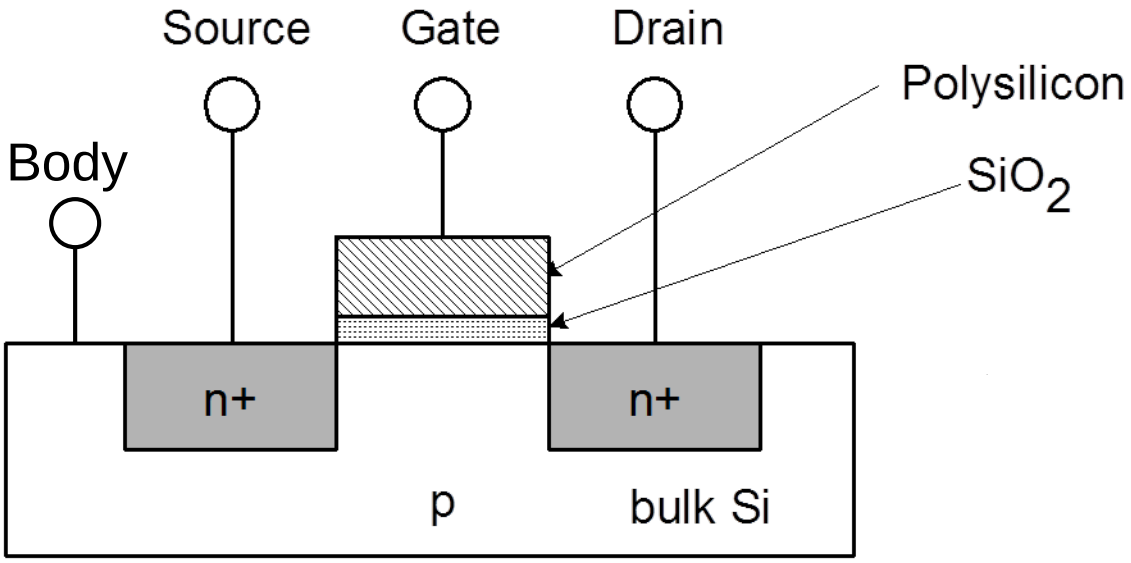


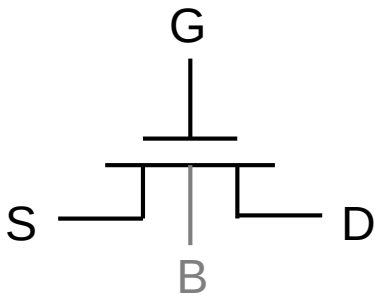
anode cathode



nMOS Transistor

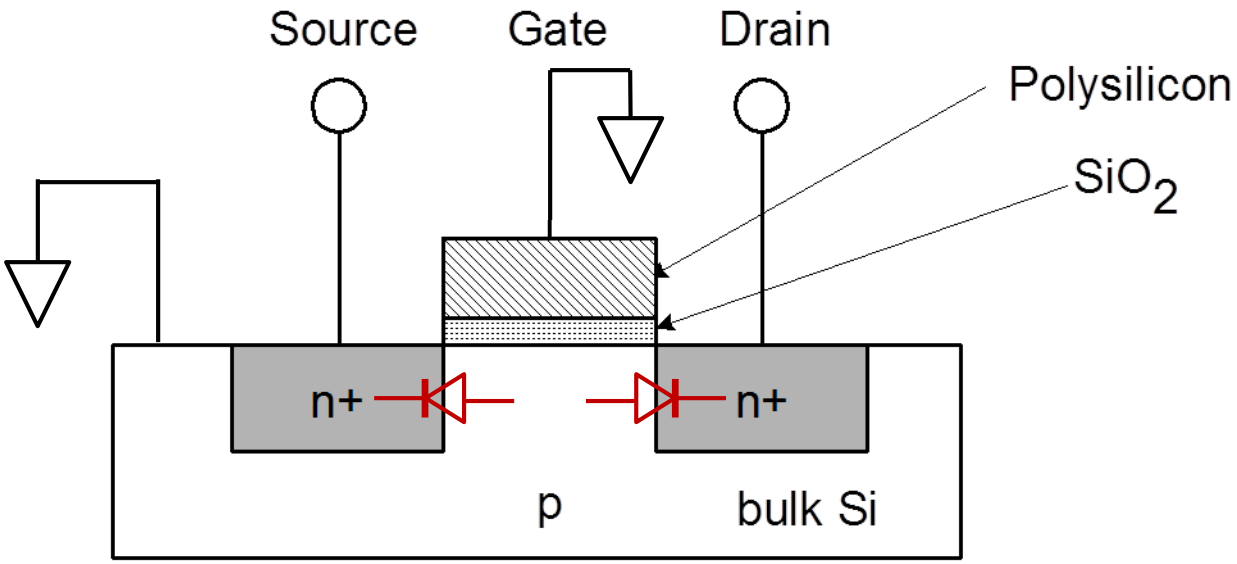
- Four terminals: gate, source, drain, body
- Gate – oxide – body stack looks like a capacitor
  - Gate and body are conductors
  - SiO_2 (oxide) is a very good insulator
  - Called metal – oxide – semiconductor (MOS) capacitor
    - Even though gate is no longer made of metal

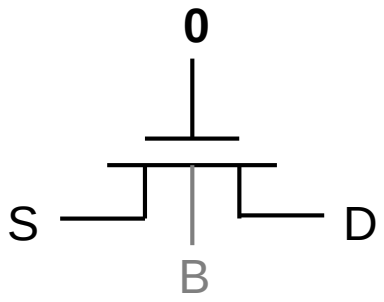




nMOS Operation (1)

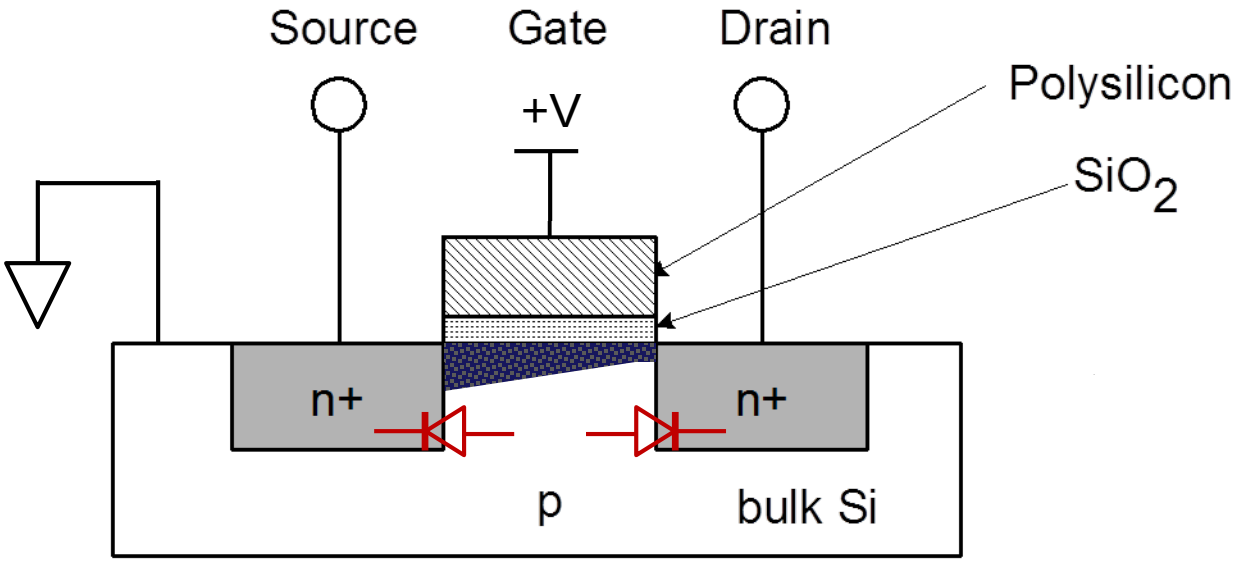
- P-type body is commonly tied to negative rail (0 V)
- When the gate is at a low voltage (~ 0V):
- Source and drain are at some positive voltage (≥0V)
  - Source (by definition) is at lower voltage than the drain
- Source-body and drain-body diodes are OFF
- No current flows, transistor is OFF

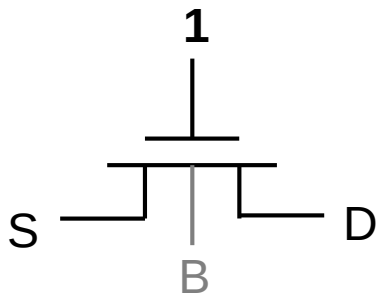




nMOS Operation (2)

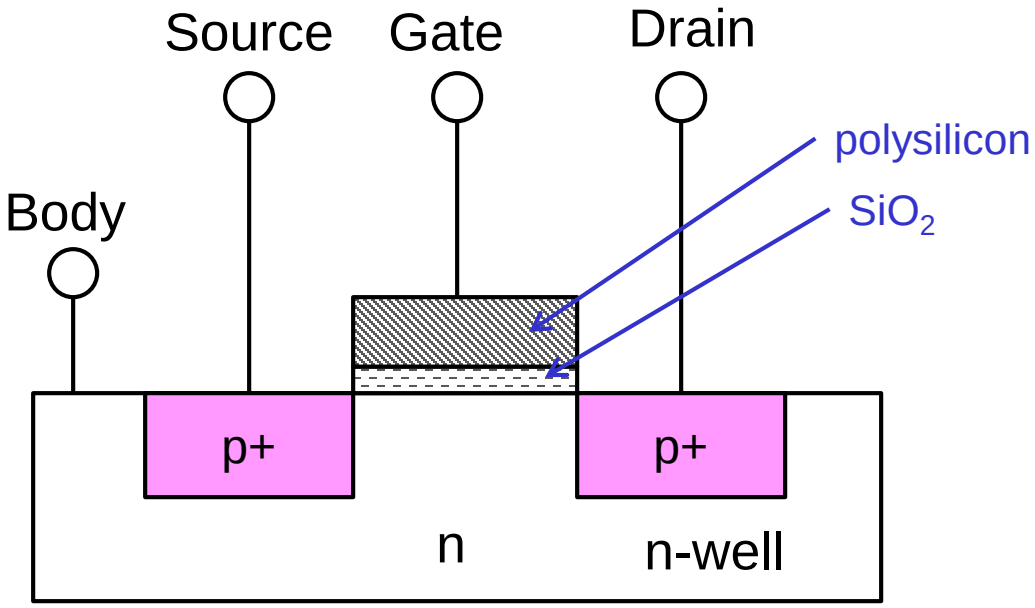
- When the gate is at a high voltage (+V volts):
- Positive charge on gate of MOS capacitor
- Negative charge (electrons) drawn out of source/drain
  - Inverts a channel under gate to n-type
- Now electrons can flow through n-type silicon from source through channel to drain, transistor is ON

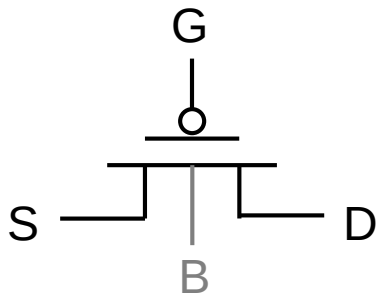




pMOS Transistor

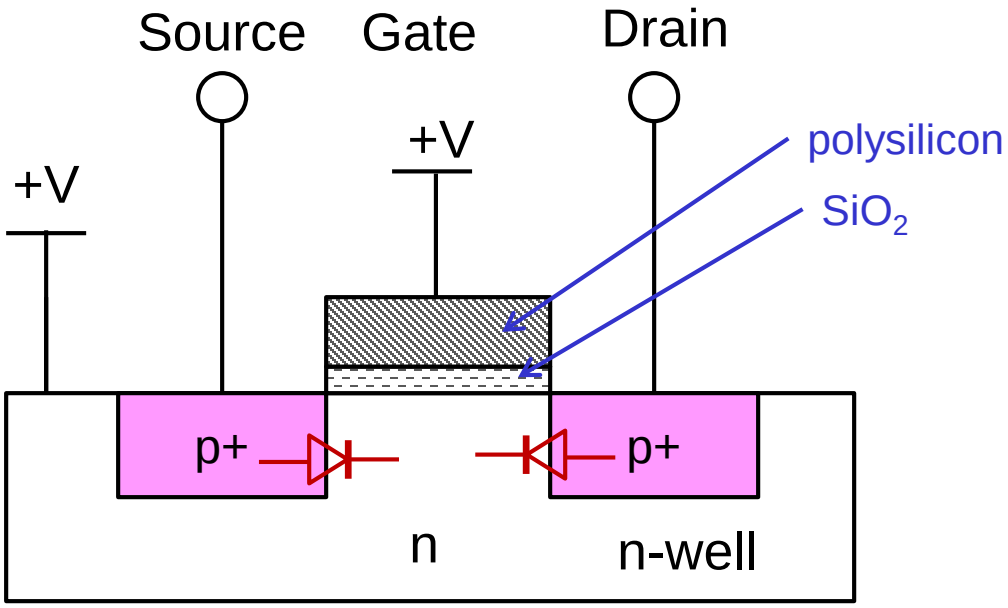
- Similar to nMOS but doping and voltages are reversed
- pMOS transistor normally built in n-well
- “Bubble” on gate of symbol indicates inverted behavior

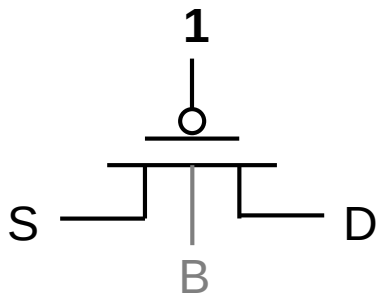




pMOS Operation (1)

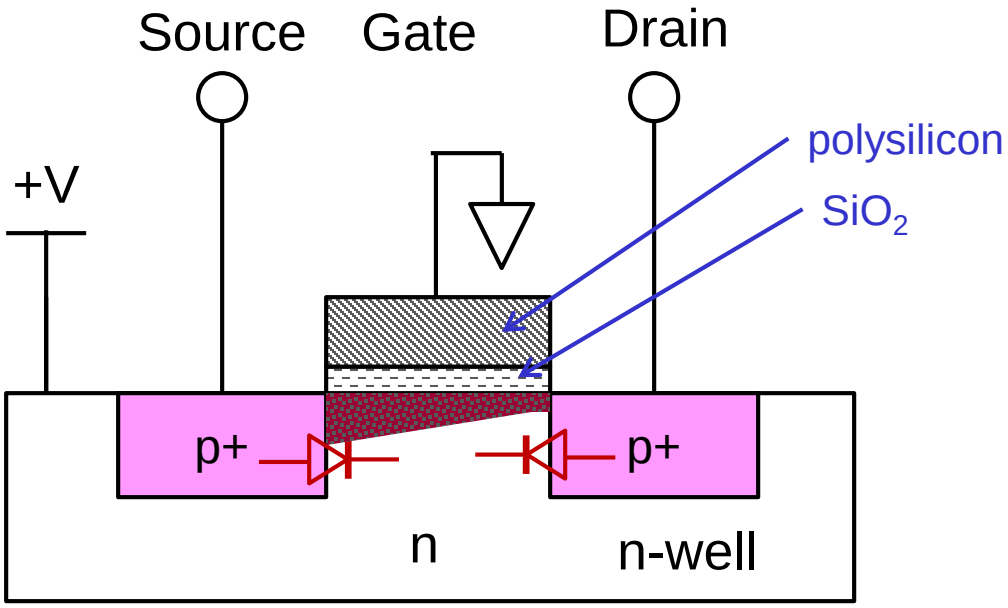
- n-type body is commonly tied to positive rail (+V)
- When the gate is also at a high voltage (~ +V),
- Source and drain are at some voltage (≤ +V)
  - source (by definition) is at higher voltage than the drain
- Source-body and drain-body diodes are OFF
- No current flows, transistor is OFF

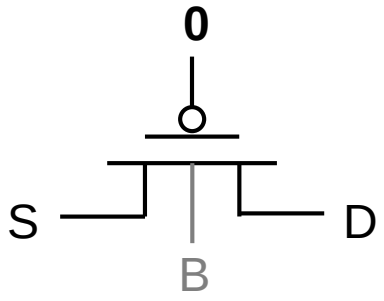




pMOS Operation (2)

- When the gate is at a low voltage (~ 0 volts):
- Negative charge on gate of MOS capacitor
- Positive charge (holes) drawn out of source/drain
- Inverts a channel under gate to p-type
- Now holes can flow through p-type silicon from source through channel to drain, transistor is ON





Power Supply Voltages

- V_{SS} ≈ negative rail ≈ GND ≈ 0 V
- In 1980's, positive rail V_{DD} = 5V
- VDD has decreased in modern processes
- High VDD would damage modern tiny transistors
  - very thin gate oxide ~ 2 nm
- Lower VDD saves power
- VDD has progressed:

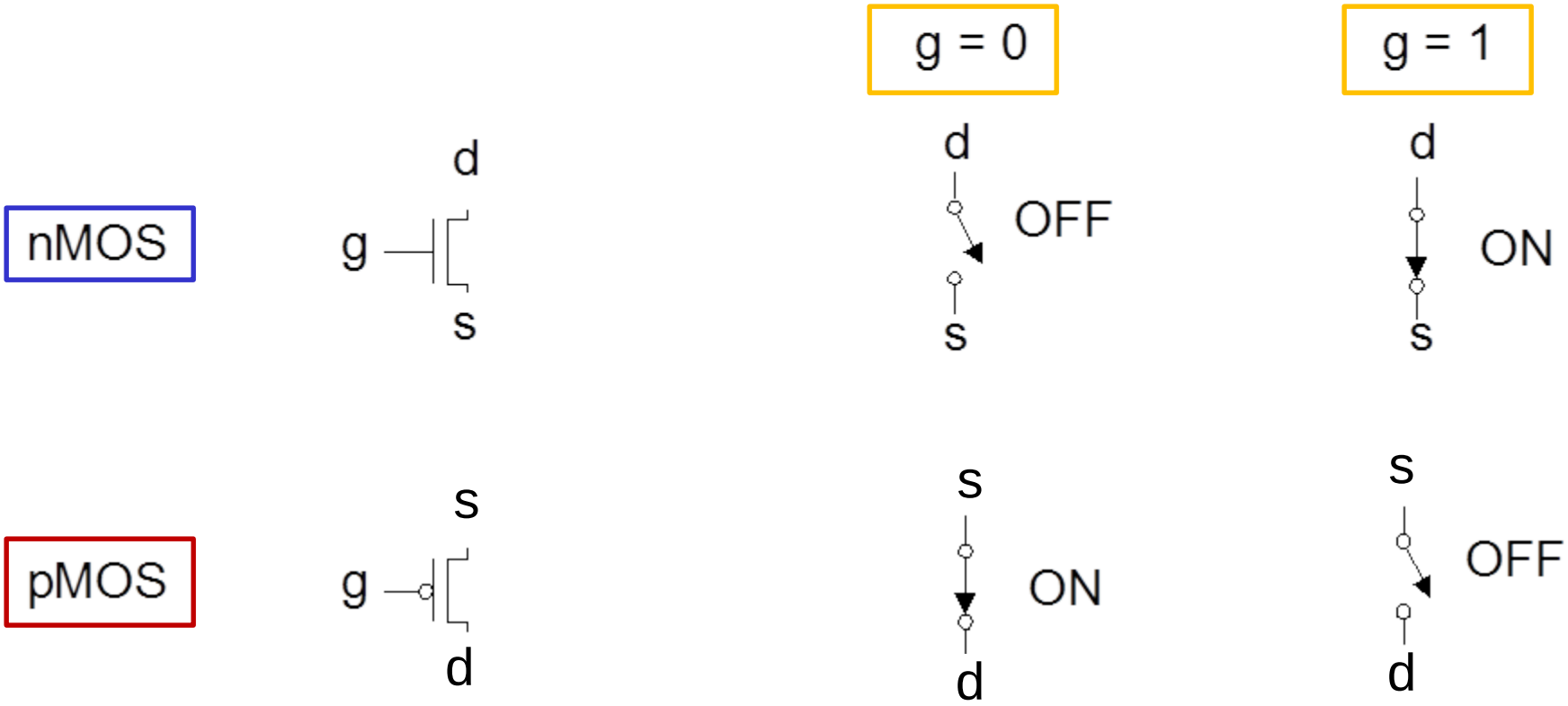
Process: 0.35μ ⇨ 0.25μ ⇨ 180nm ⇨ 130nm ⇨ 90nm ⇨ 65nm

VDD: 3.3V ⇨ 2.5V ⇨ 1.8V ⇨ 1.5V ⇨ 1.2V ⇨ 1.0V ⇨ ??

- In CMOS digital circuits, define:
  - GND ≡ logical '0'
  - VDD ≡ logical '1'

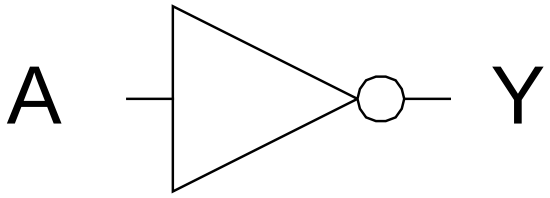
Transistors as Switches

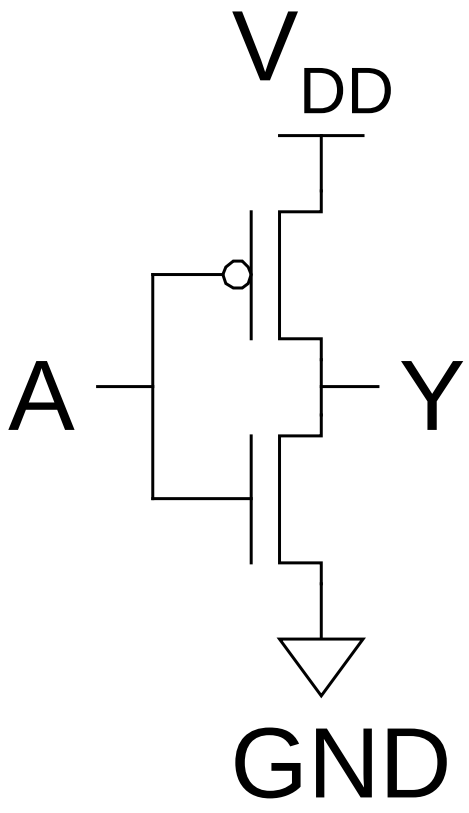
- In simplest model, we can view MOS transistors as electrically controlled switches
- Voltage at gate controls path from source to drain



CMOS Inverter

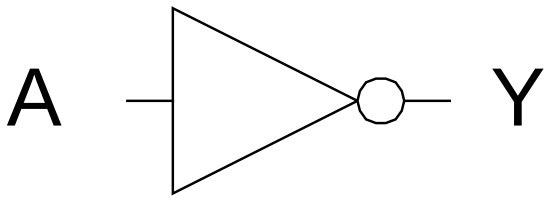
A	Y
0	
1	

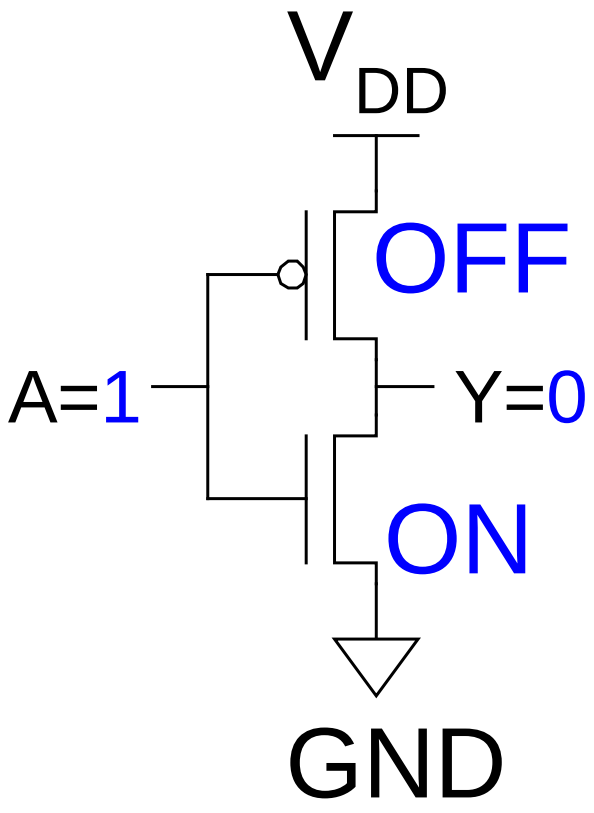




CMOS Inverter

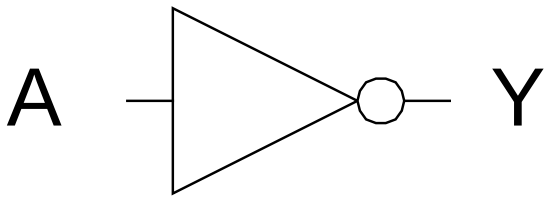
A	Y
0	
1	**0**

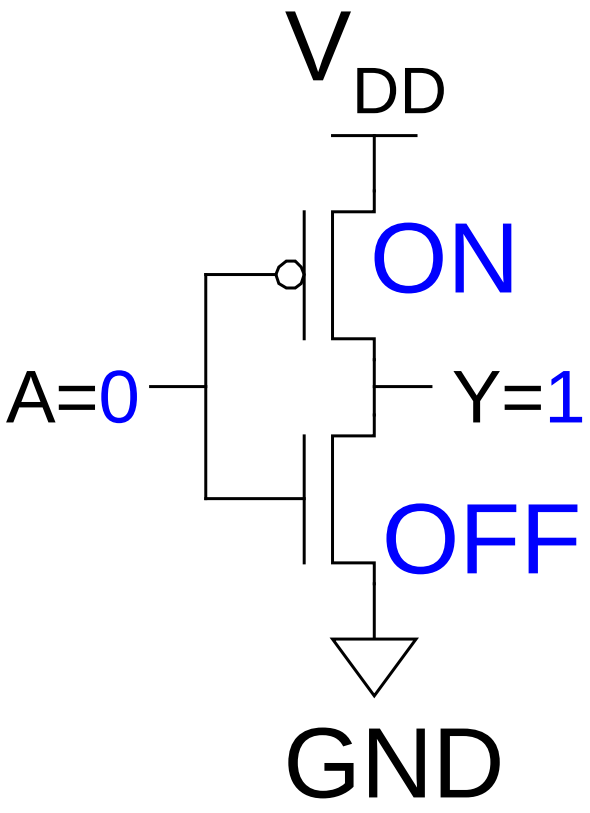




CMOS Inverter

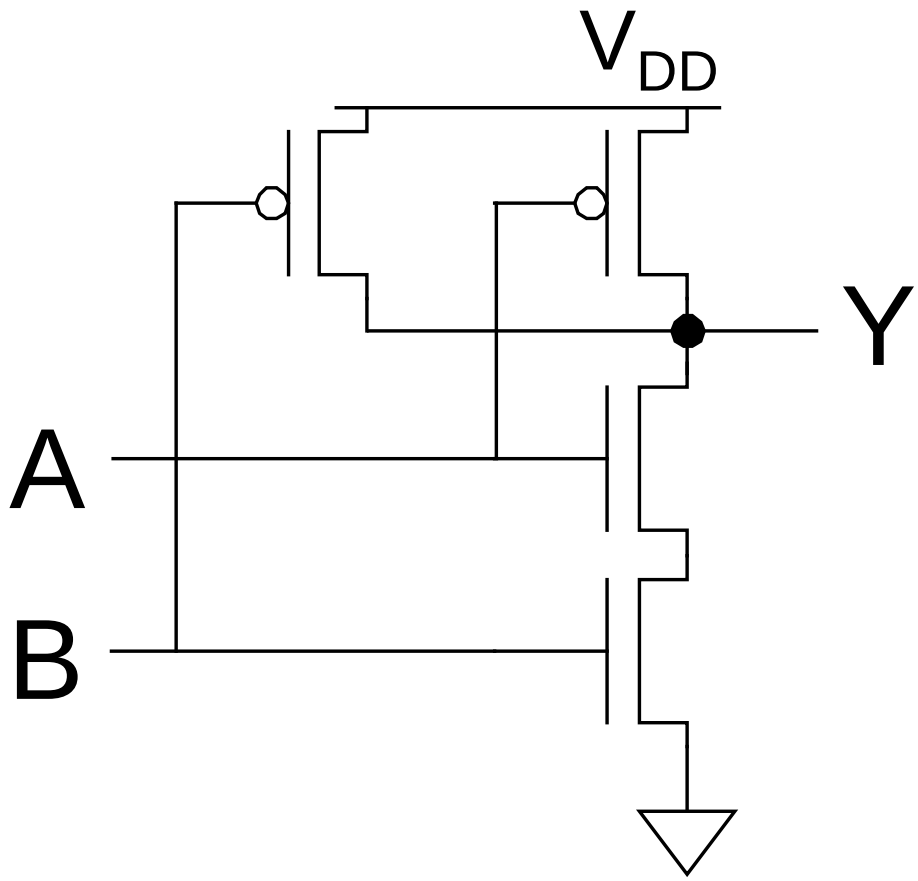
A	Y
0	**1**
1	0





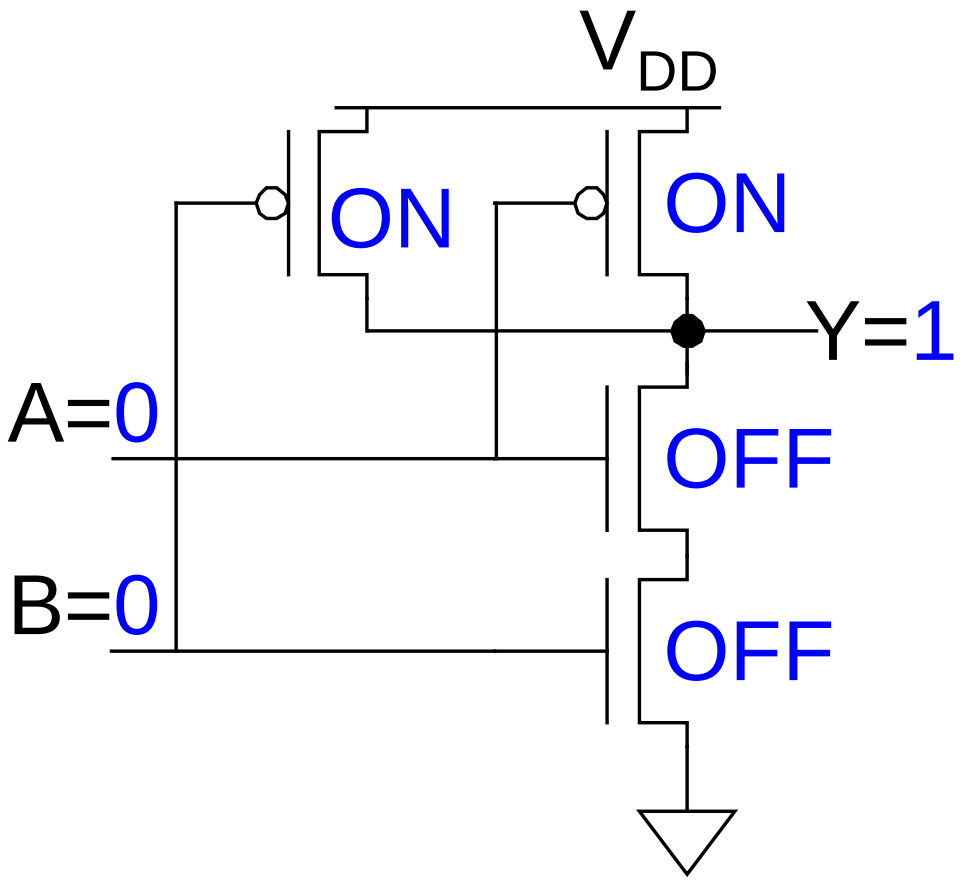
CMOS 2-input NAND Gate

A	B	Y
0	0	
0	1	
1	0	
1	1	



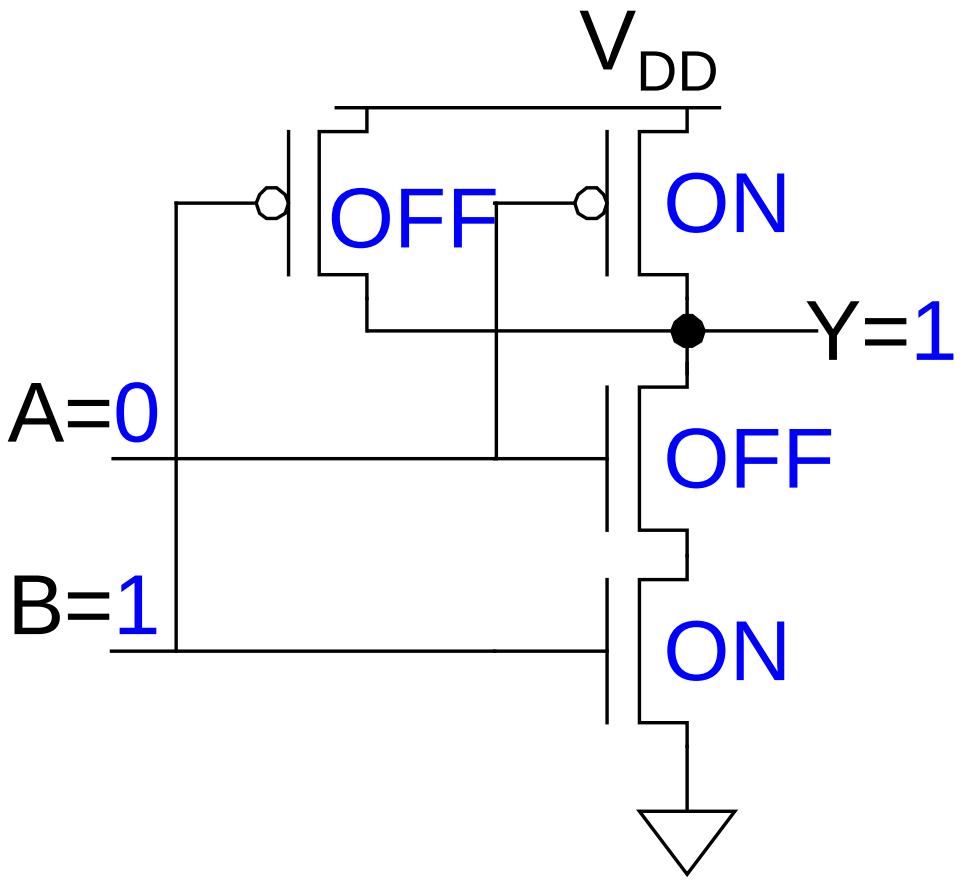
CMOS 2-input NAND Gate

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



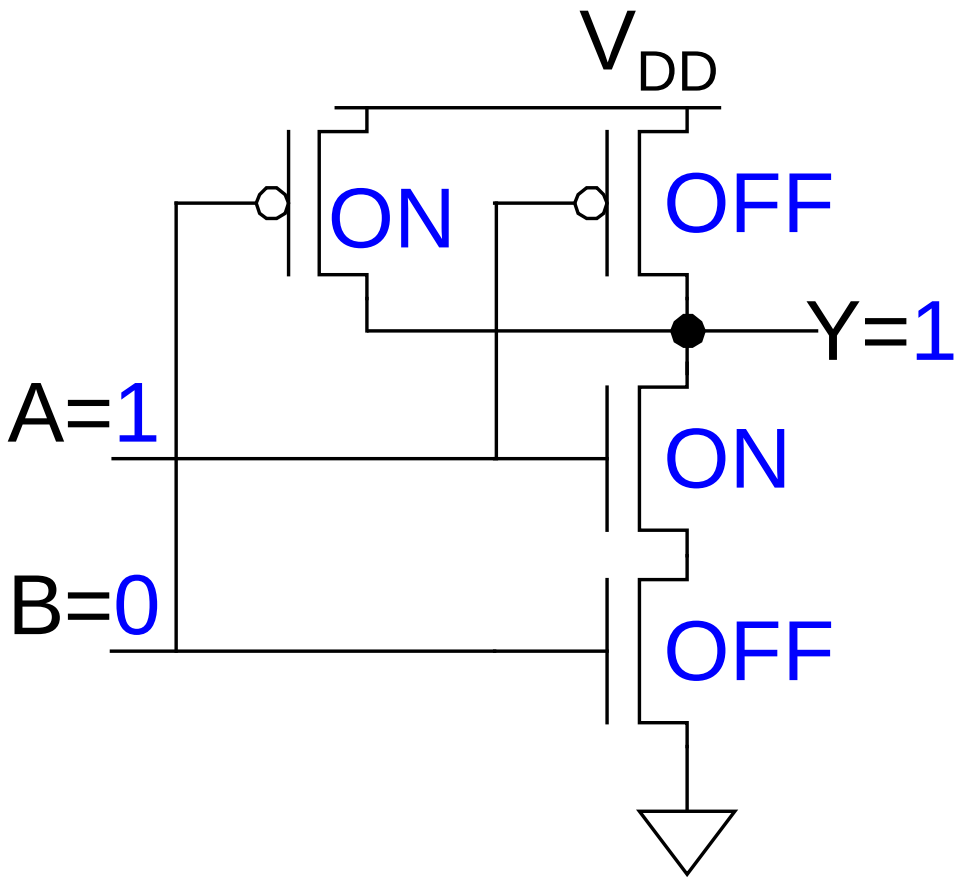
CMOS 2-input NAND Gate

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



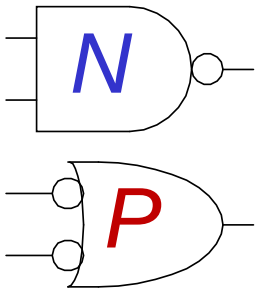
CMOS 2-input NAND Gate

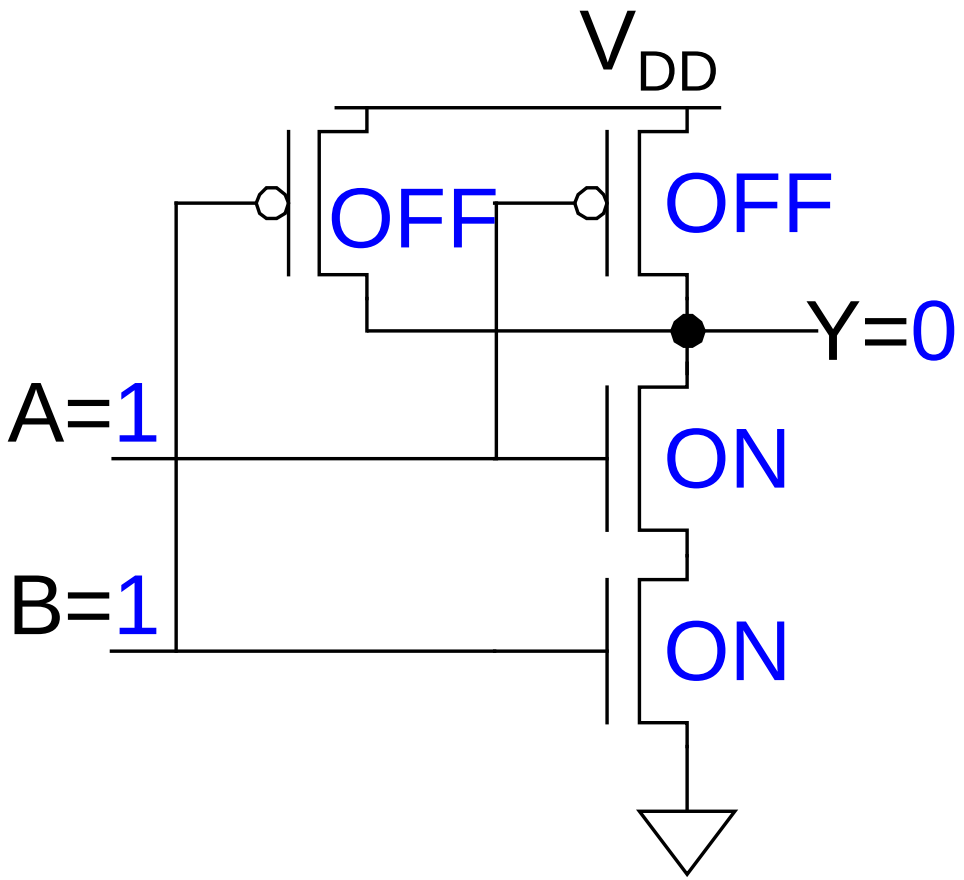
A	B	Y
0	0	1
0	1	1
1	**0**	**1**
1	1	



CMOS 2-input NAND Gate

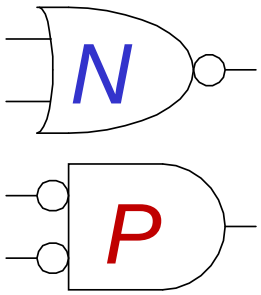
A	B	Y
0	0	1
0	1	1
1	0	1
1	**1**	**0**

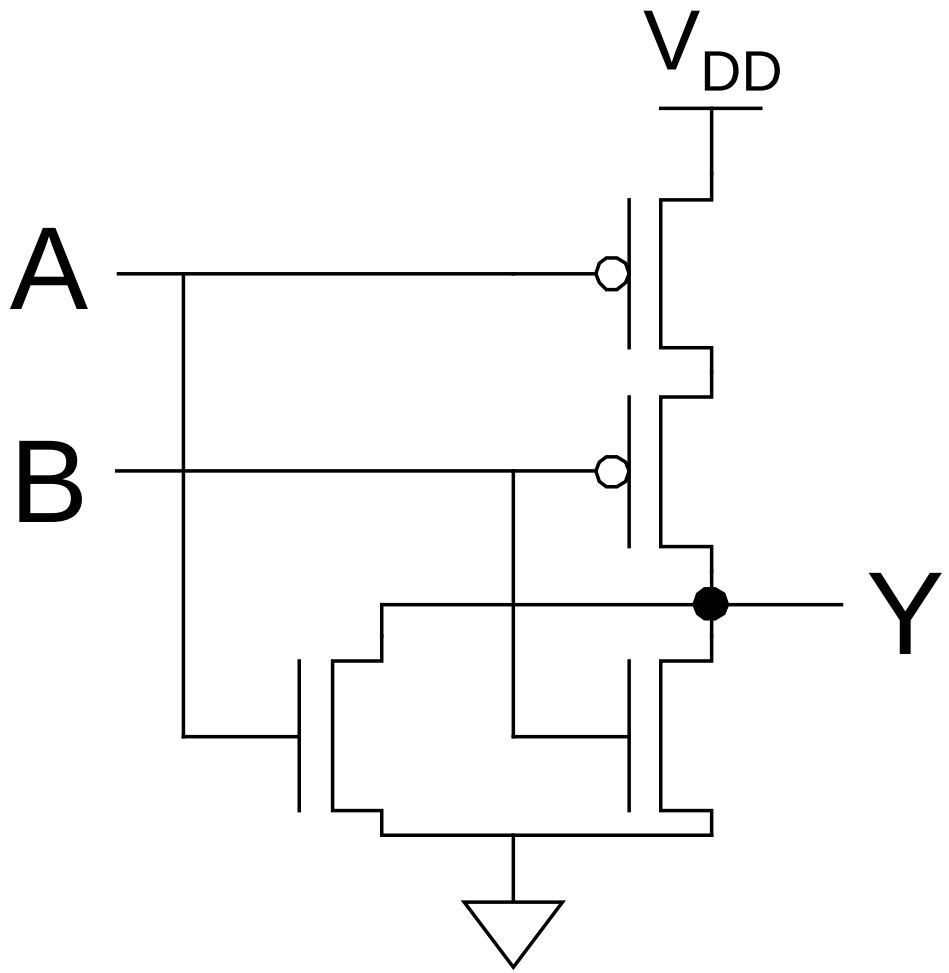




CMOS 2-input NOR Gate

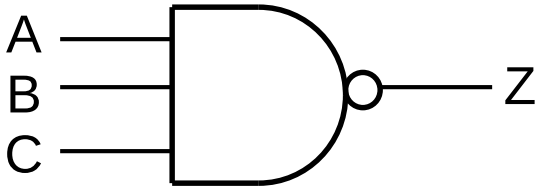
A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0





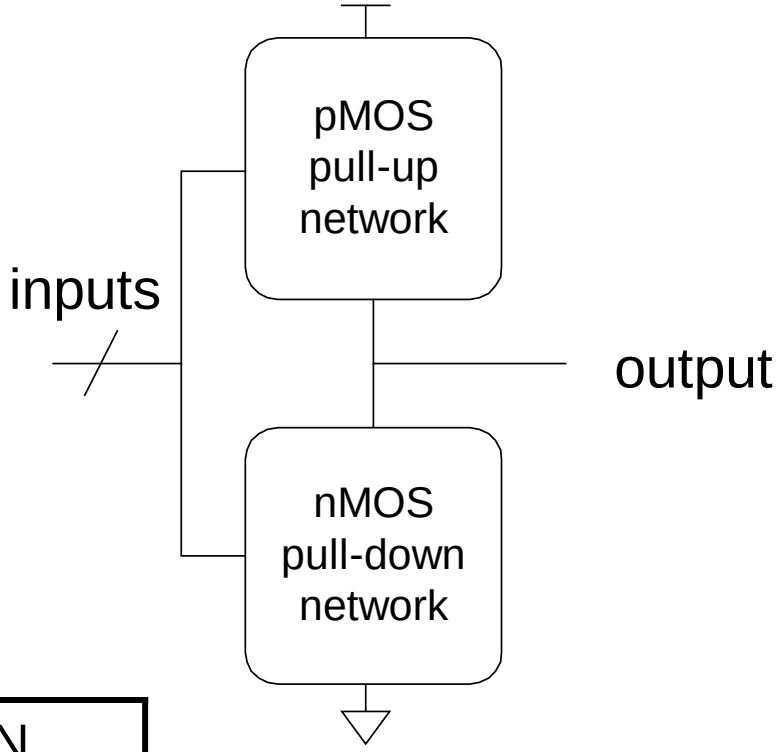
CMOS Gate Design

- Draw the transistor level schematic of a 3-input CMOS NAND gate:



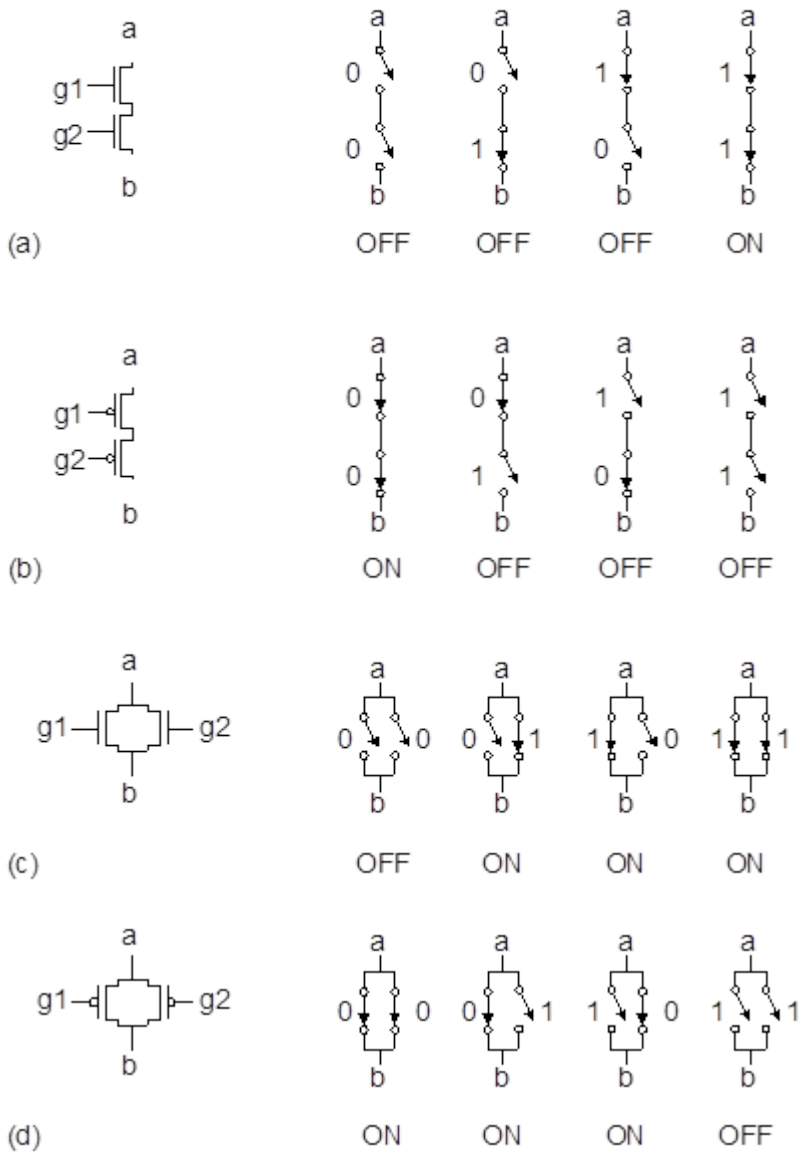
Complementary CMOS Gates

- nMOS pull-down network
- pMOS pull-up network
- static combinational CMOS logic



	Pull-up OFF	Pull-up ON
Pull-down OFF	Z (float)	**1**
Pull-down ON	**0**	X (crowbar)

Series & Parallel Conduction Paths

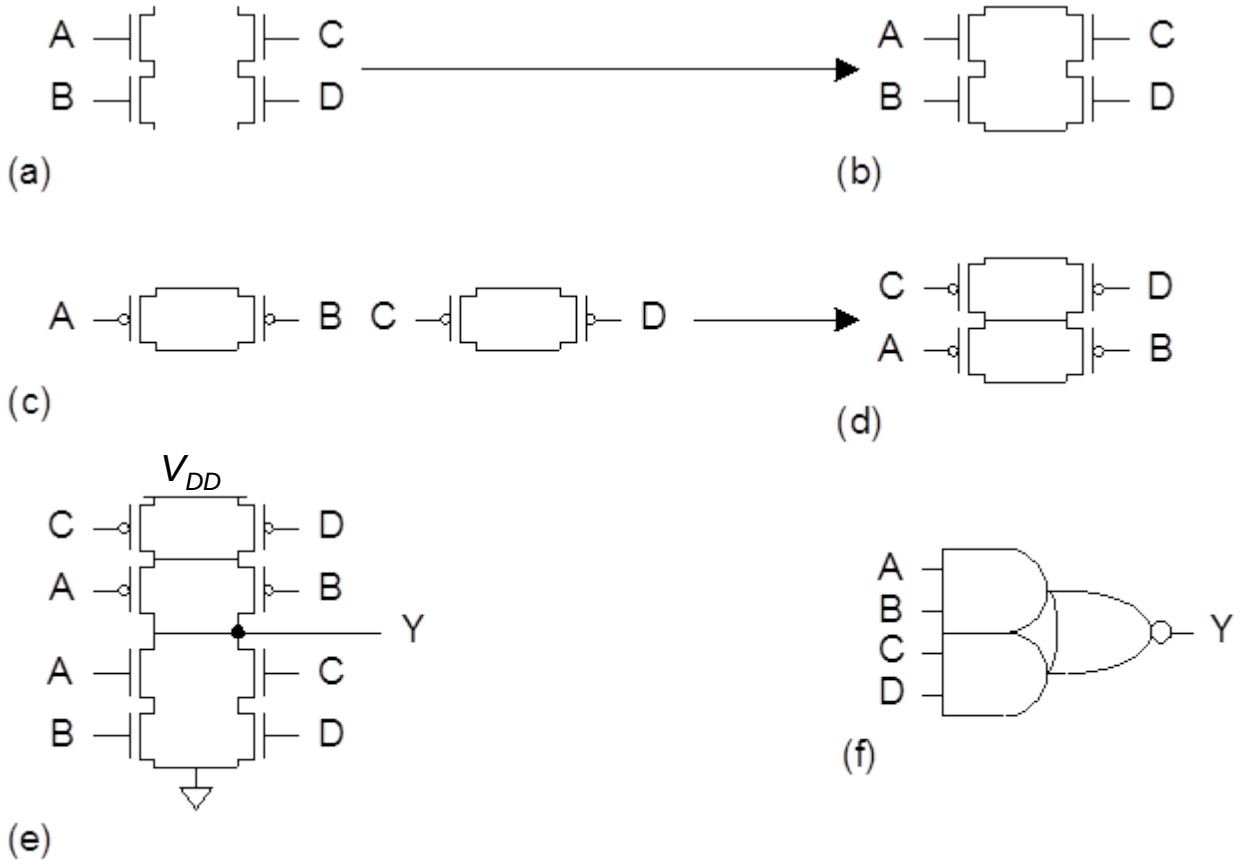


- nMOS: 1 = ON
- pMOS: 0 = ON
- Series: both must be ON
- Parallel: either can be ON

- To ensure that gate is always driven to 0 or 1:
- Pull-up network must be topological complement of pull-down network
  - parallel ⇨ series
  - series ⇨ parallel

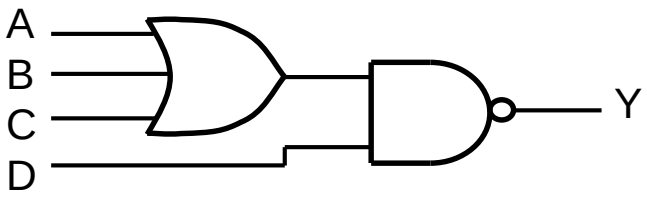
Compound Gates

- We can generate any inverting combinational function with a network of series and parallel nMOS transistors and a complementary network of pMOS transistors
- e.g., $Y = \overline{A.B + C.D}$ *and-or-invert gate*: AOI22



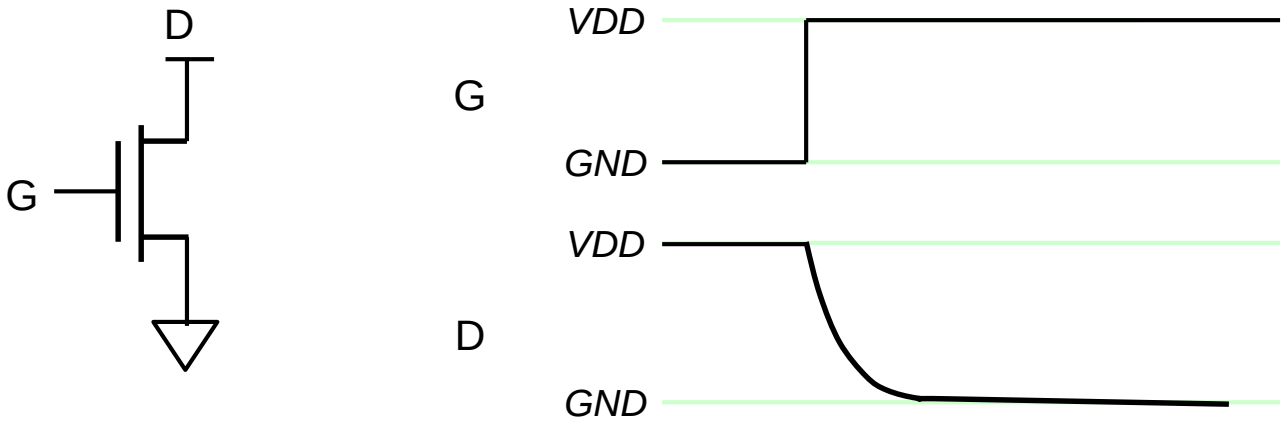
Example: O3AI

- $Y = \overline{(A + B + C).D}$



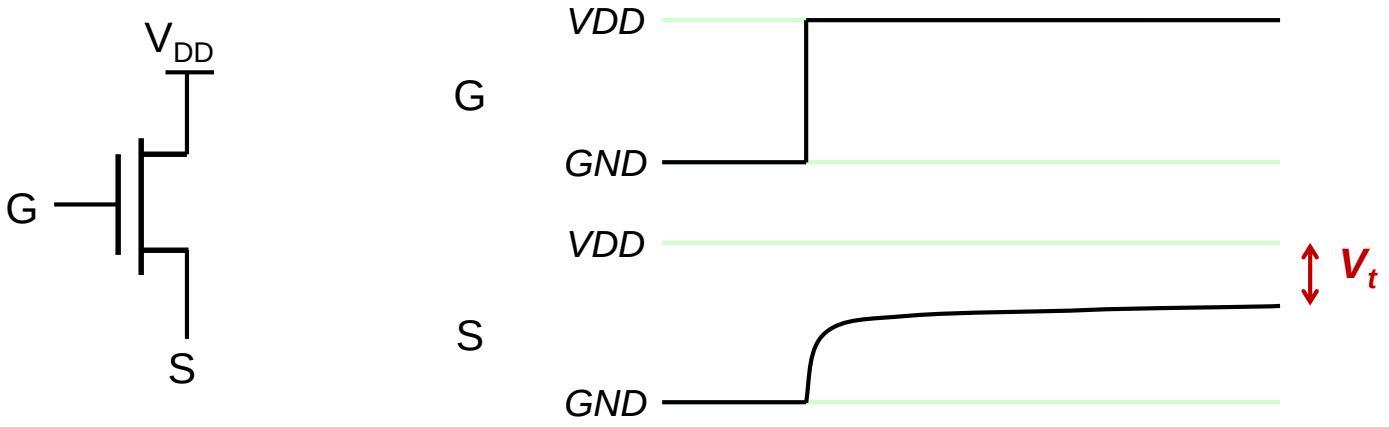
Signal Strength

- In a compound gate, nMOS transistors are always used to pull down to GND and pMOS are always used to pull up to V_{DD}



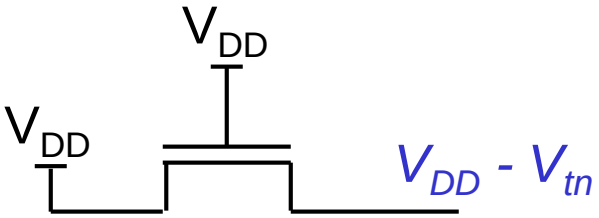
- Once gate goes high, $V_g - V_s = V_{DD} > V_{th}$ (threshold voltage)
- Transistor stays on as drain is pulled all way down to GND
- Could we use an nMOS transistor to pull-up to V_{DD}?

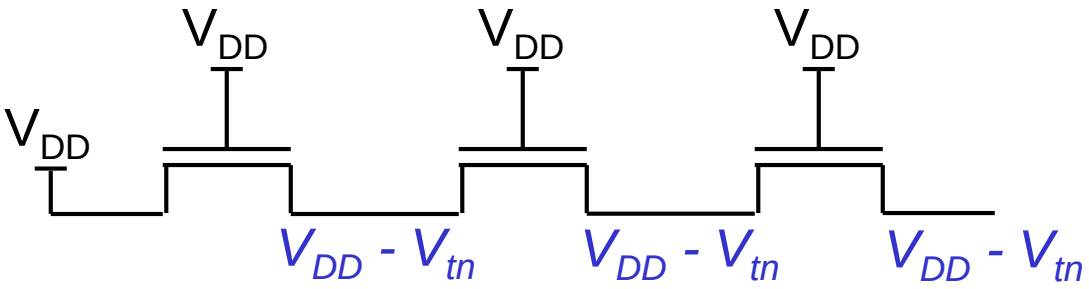
Pulling up with an nMOS

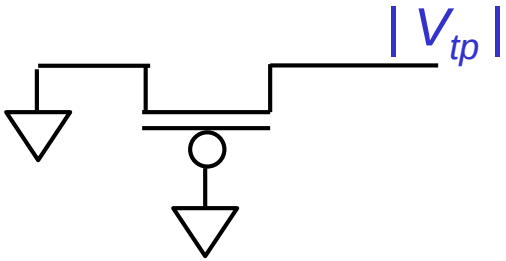


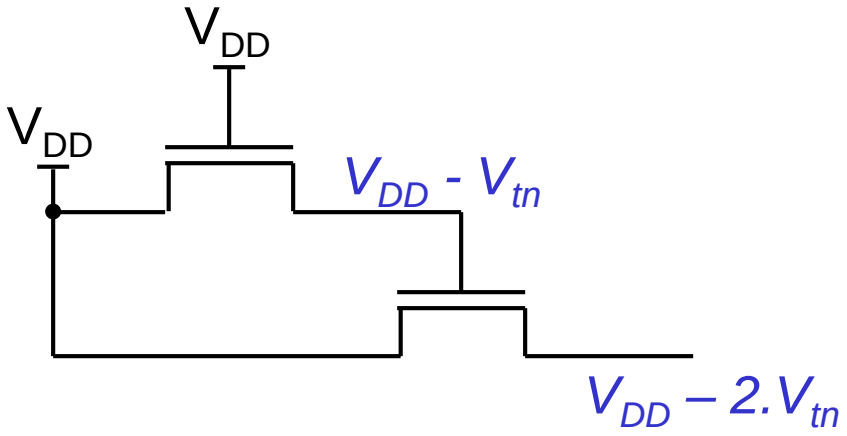
- In this configuration, source voltage is changing
- As $V_g - V_s$ approaches V_{th}, transistor starts to turn off
- Weak conduction leads to degraded final value
  - never reaches V_{DD}. V_s asymptotes towards $V_{DD} - V_t$
- As a switch, we say nMOS drives (passes) a strong 0 but a degraded or weak 1
  - Similarly pMOS drives a strong 1 but a degraded or weak 0

Cascaded Pass Transistors



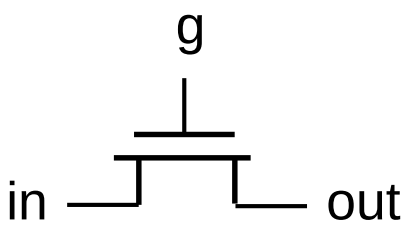


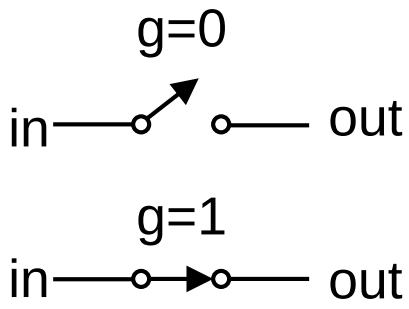


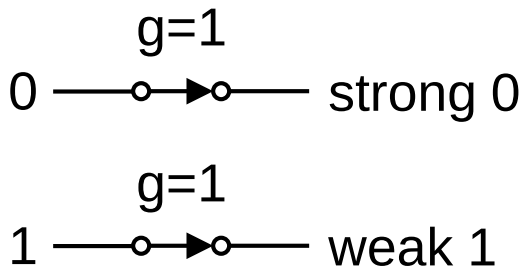


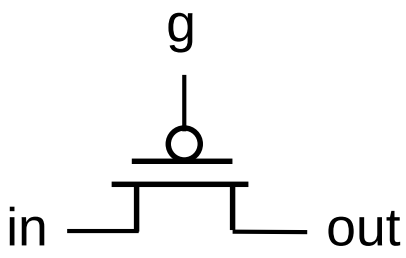
Pass Transistors

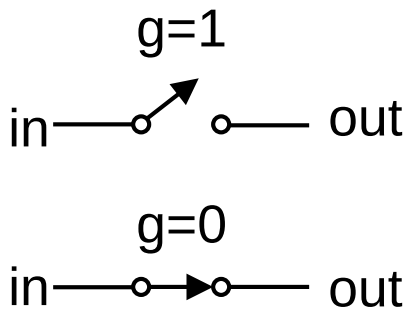
- So far, we have used nMOS to switch (drive) output to GND and pMOS to switch (drive) output to VDD in response to various input signals
- We can also use MOS transistors to switch the input signals themselves

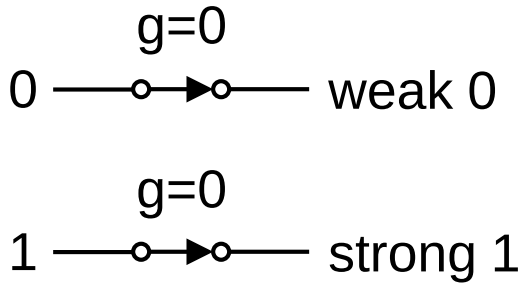






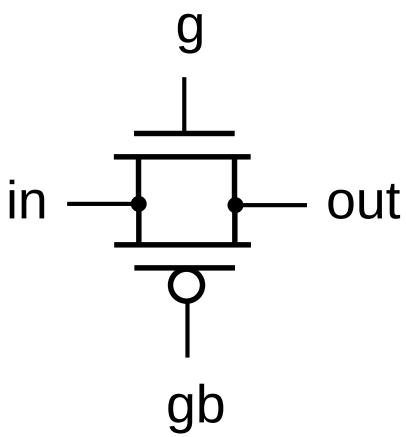


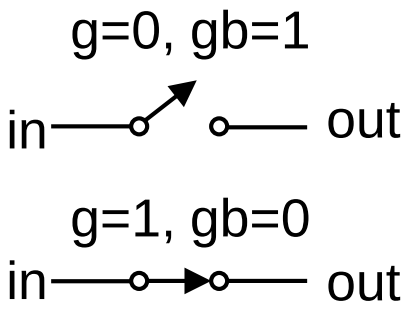


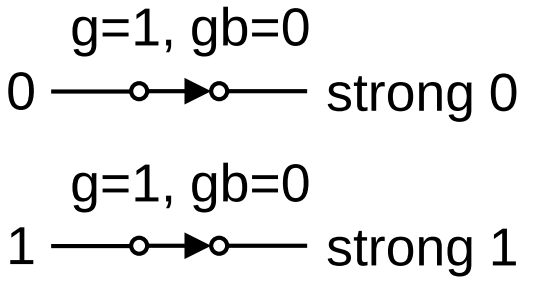


Transmission Gate

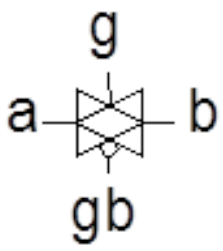
- Transmission gate is a pMOS and nMOS pass transistor in parallel
- Passes a strong 0 and a strong 1

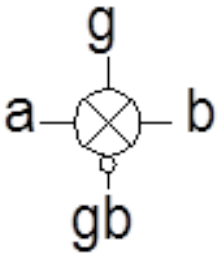


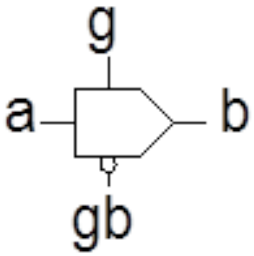




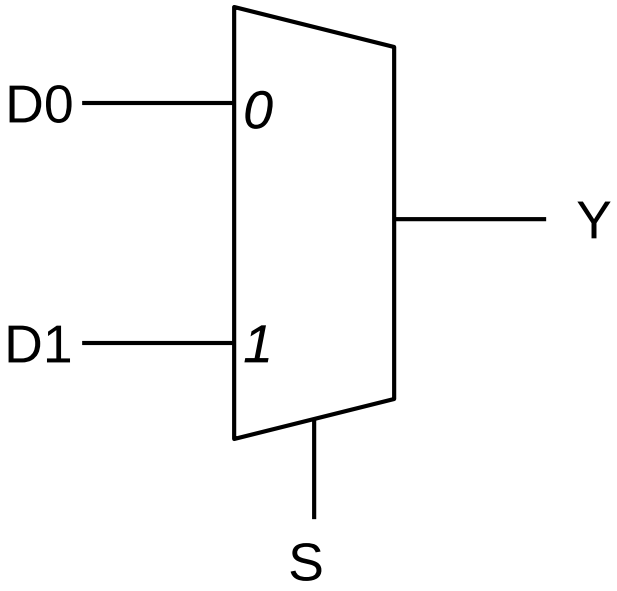
- Common schematic symbols:







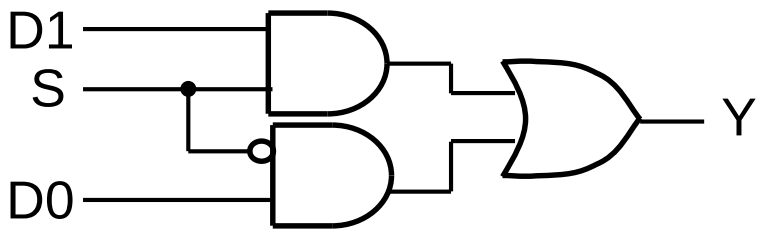
2:1 Multiplexer



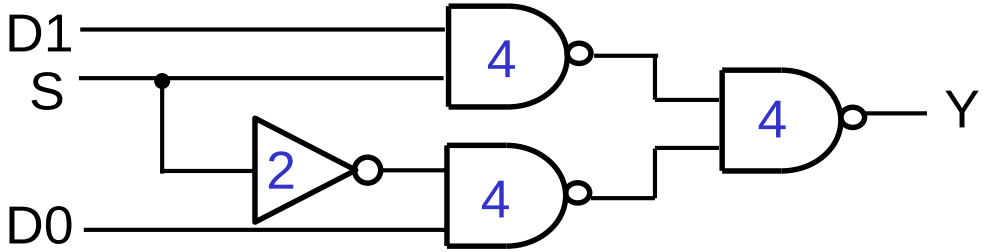
S	D1	D0	Y
0	X	0	0
0	X	1	1
1	0	X	0
1	1	X	1

Mux Design using Standard Logic Gates

- $Y = \overline{S}.D0 + S.D1$

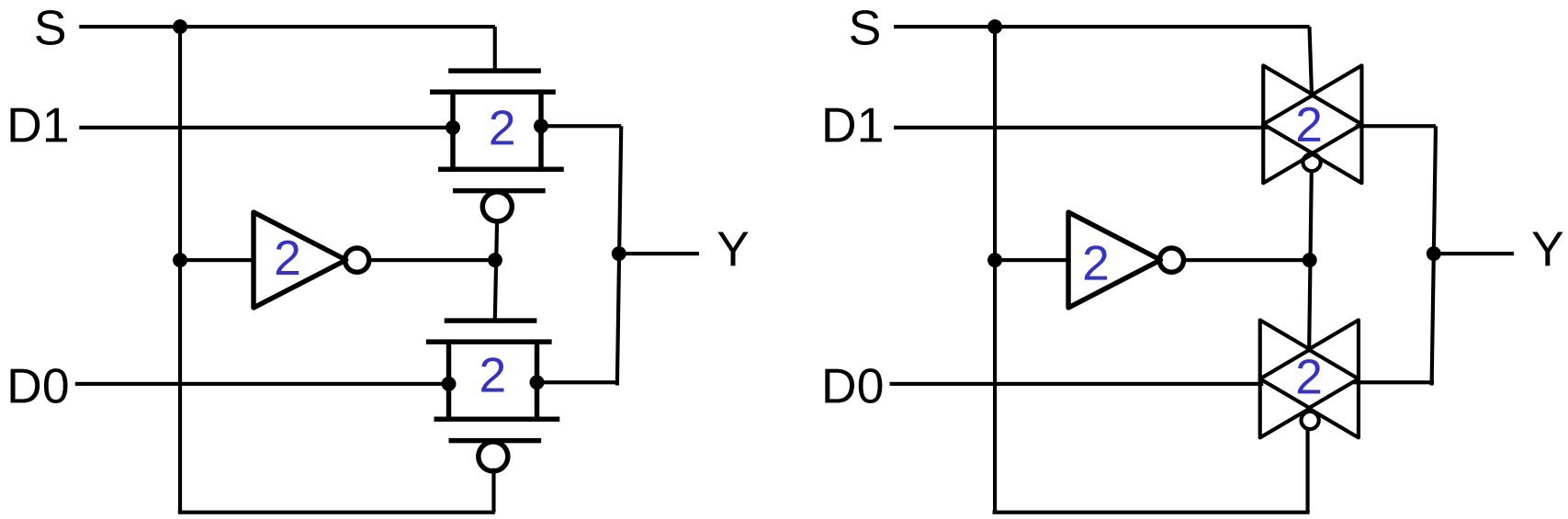


using complimentary inverting gates



- Requires 14 transistors

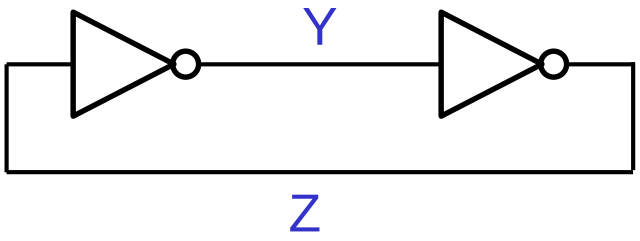
Mux Design using Transmission Gates



- Requires only 6 transistors
- Use with caution: non-restored logic
- Long chains of transmission gates lead to long delays and degraded levels

Storage Elements

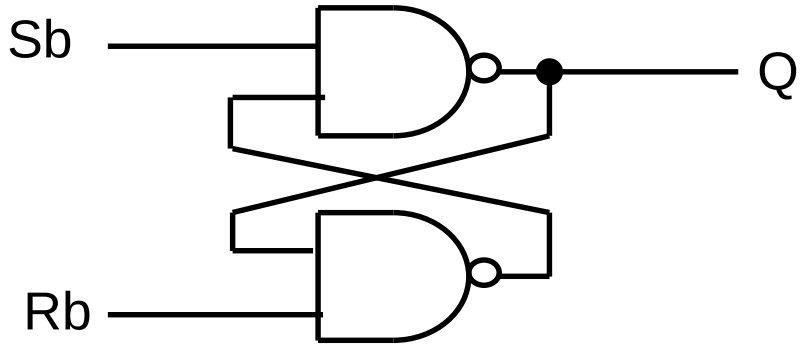
- Basic static storage element is cross-coupled inverter



- Positive feedback drives circuit into one of two stable states
- Either: (Y=1, Z=0) *OR* (Y=0, Z=1)
- Circuit will hold state indefinitely
  - restoring effect of digital logic eliminates degradation of stored levels over time
- How do we change the state?

RS Latch

- Simple “writable” storage element



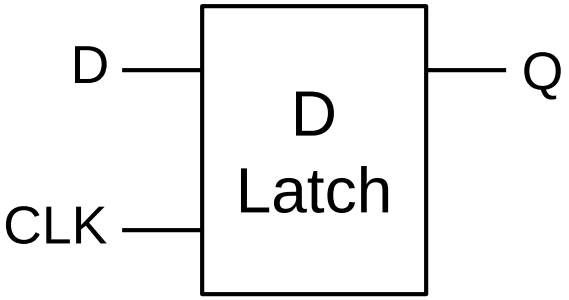
Rb	Sb	Q
0	1	0
1	0	1
1	1	*no change*
0	0	*illegal*

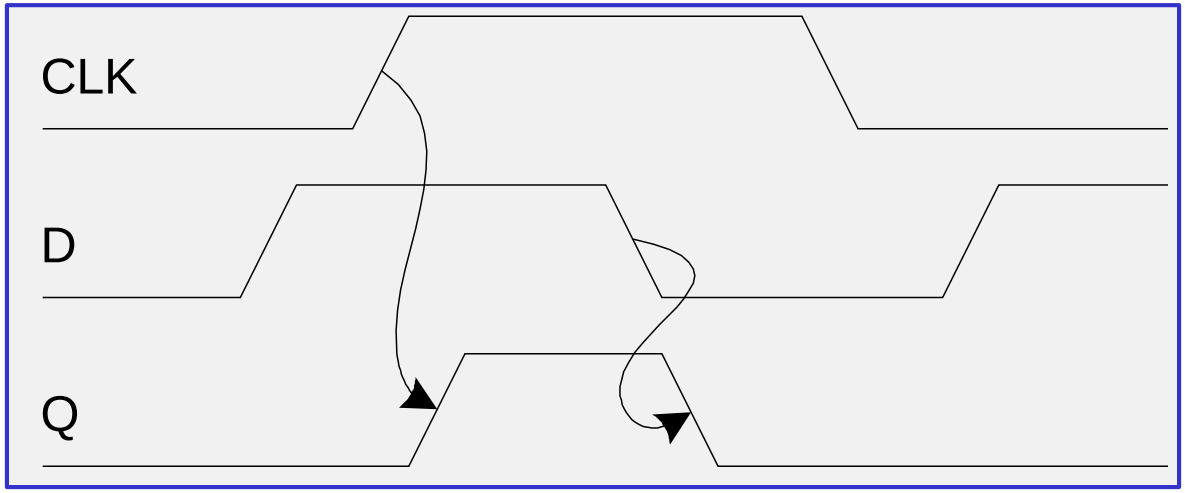
- Normally, Sb and Rb are both 1
- When Sb=0, Q is set to 1
- When Rb=0, Q is reset to 0

D Latch

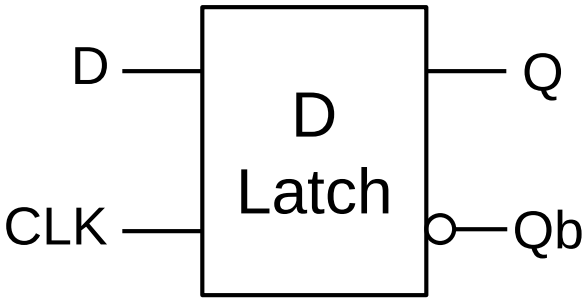
- When CLK = 1, latch is transparent
- D flows through to Q like a buffer
- When CLK = 0, the latch is opaque
- Q holds its old value independent of D
- a.k.a. transparent latch or level-sensitive latch

D	CLK	Q
0	1	0
1	1	1
0	0	*no change*
1	0	*no change*

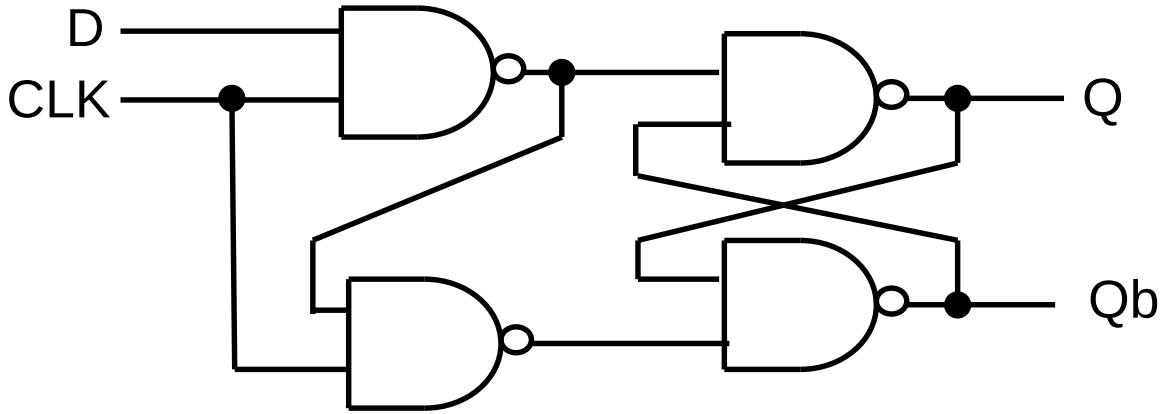




D Latch using Standard Logic Gates

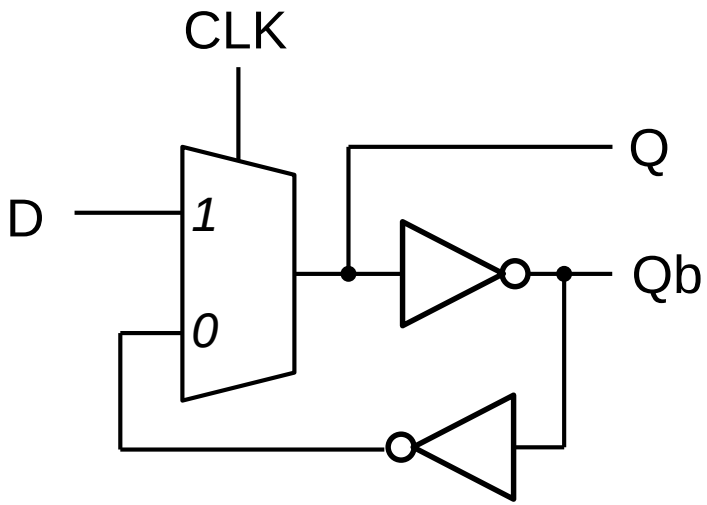


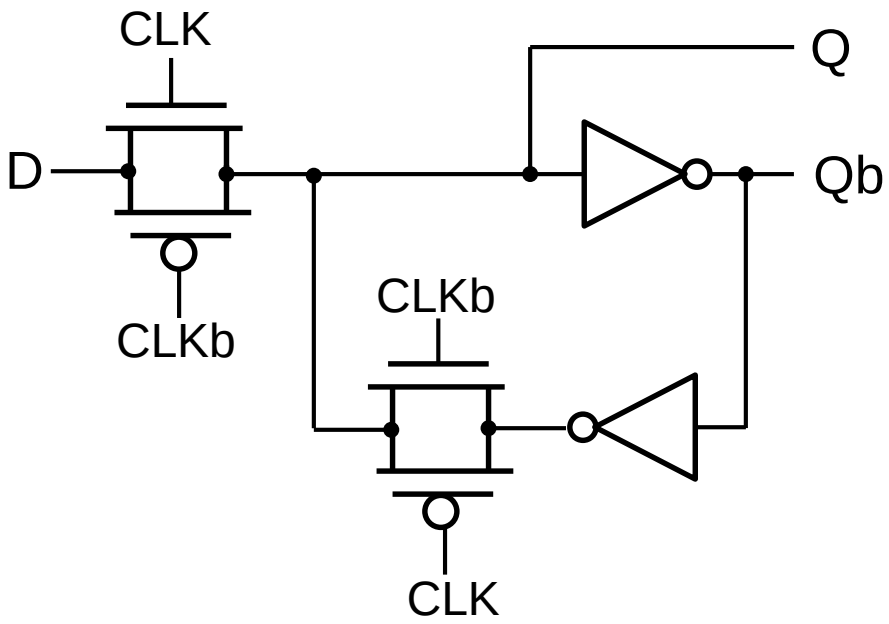
D	CLK	Q	Qb
0	1	0	1
1	1	1	0
0	0	*no change*	*no change*
1	0	*no change*	*no change*



- Uses 16 transistors
- Up to 4 gate delays (D to Q)

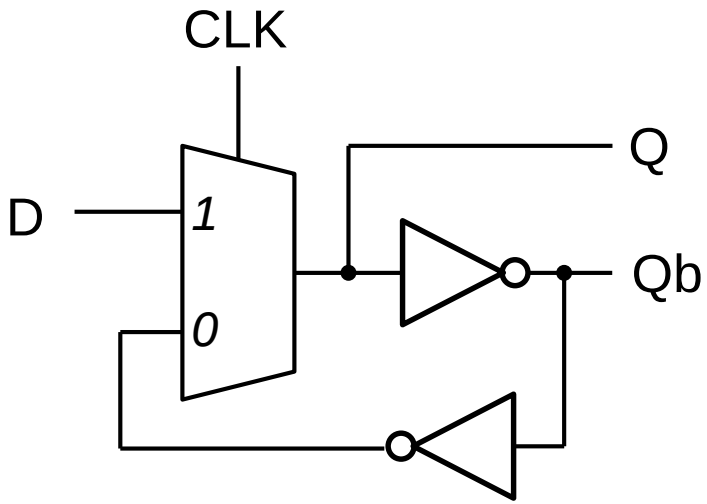
D Latch using Transmission Gate

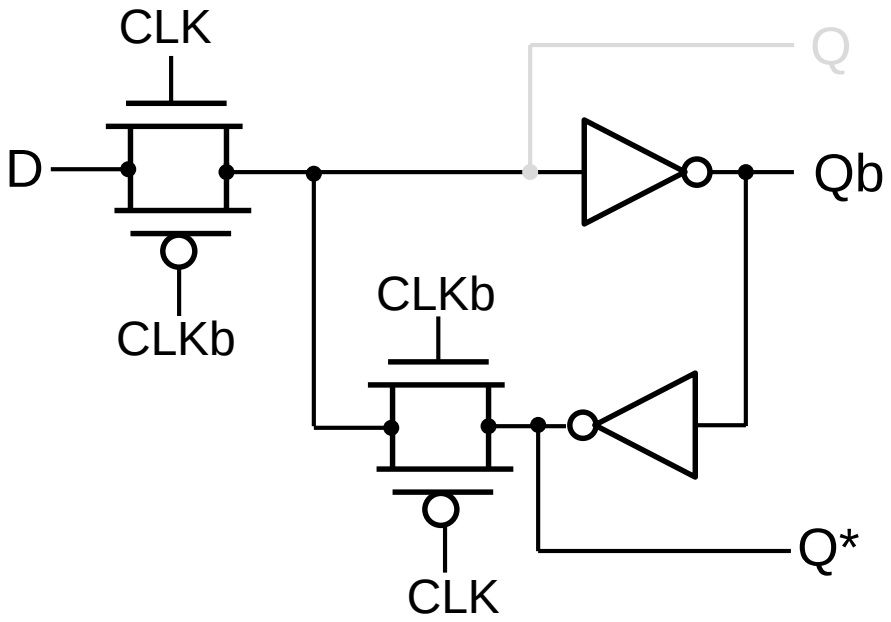




- Multiplexer chooses D or stored Q
- Uses 8 (+2) transistors
- Fast response D to Q
- Q is non-restored

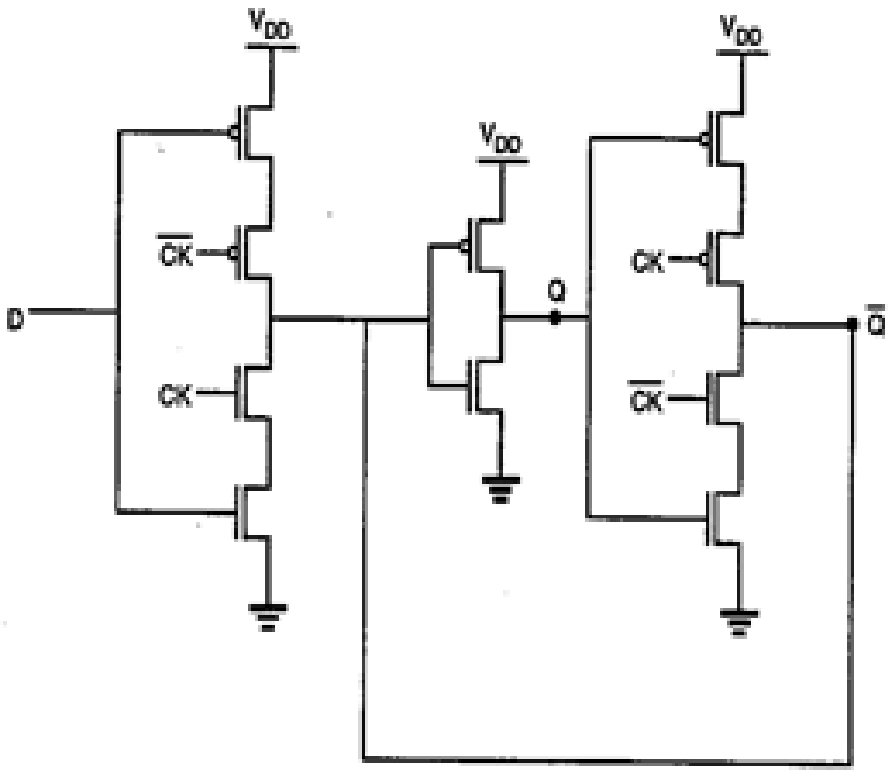
D Latch using Transmission Gate





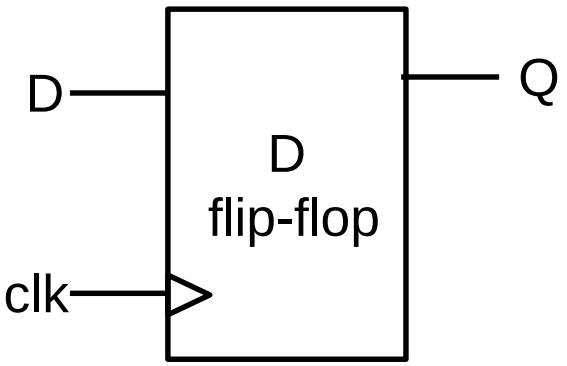
- Multiplexer chooses D or stored Q
- Uses 8 (+2) transistors
- Fast response D to Q
- Q is non-restored
- Q* is slower response, but fully restored

Alternative CMOS D Latch



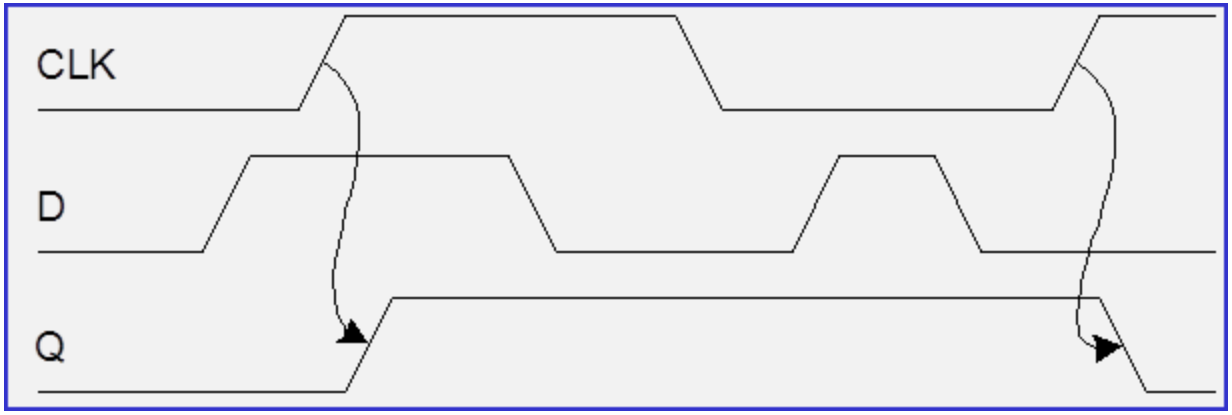
- What is happening here?

D Flip-flop



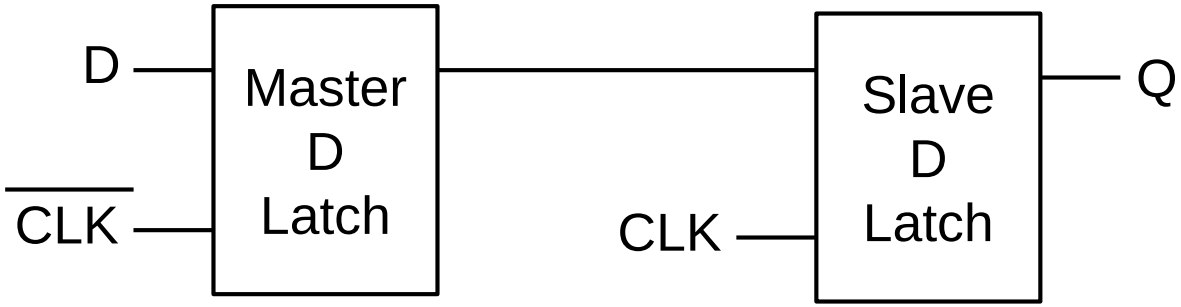
clk	D	Q
0	X	no change
1	X	no change
↑	1	1
↑	0	0

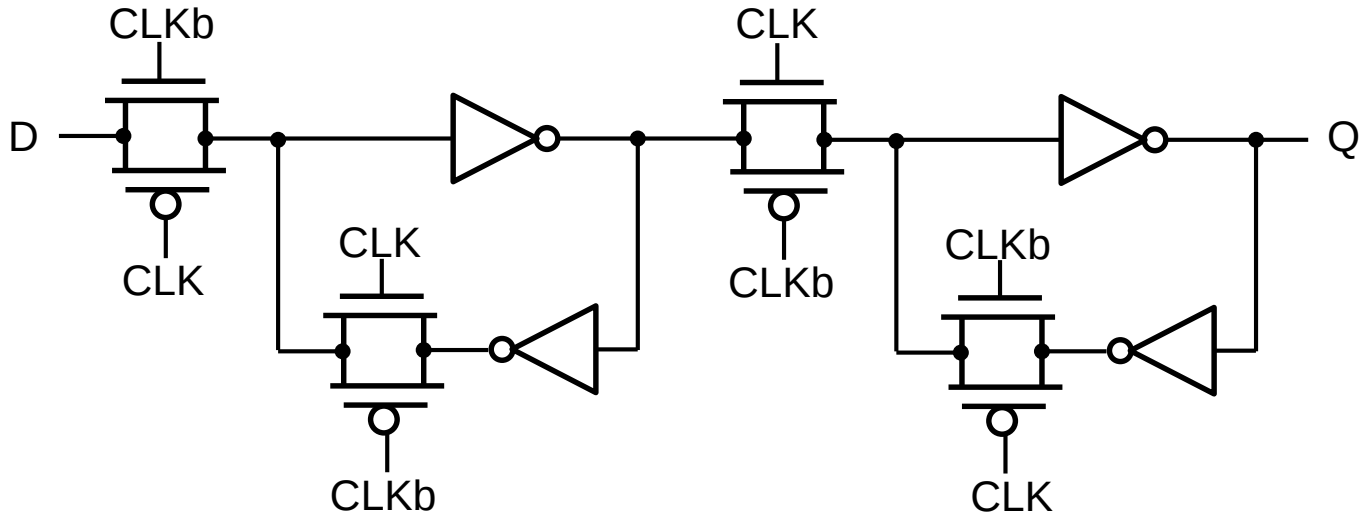
- When CLK rises, D is copied to Q
- At all other times, Q holds its value
- a.k.a. edge-triggered flip-flop, master-slave flip-flop



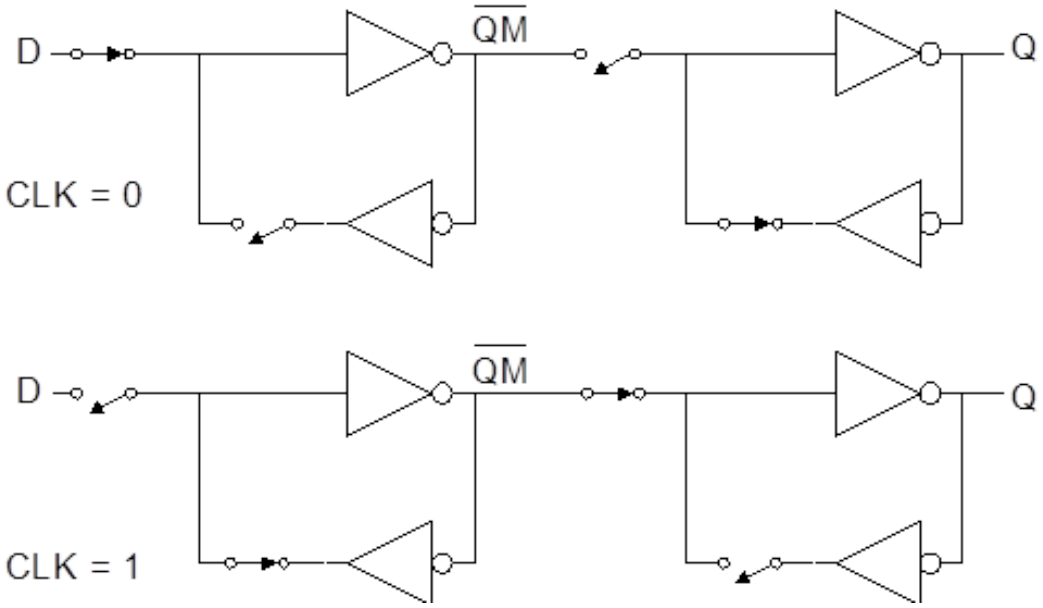
Master-Slave Latches

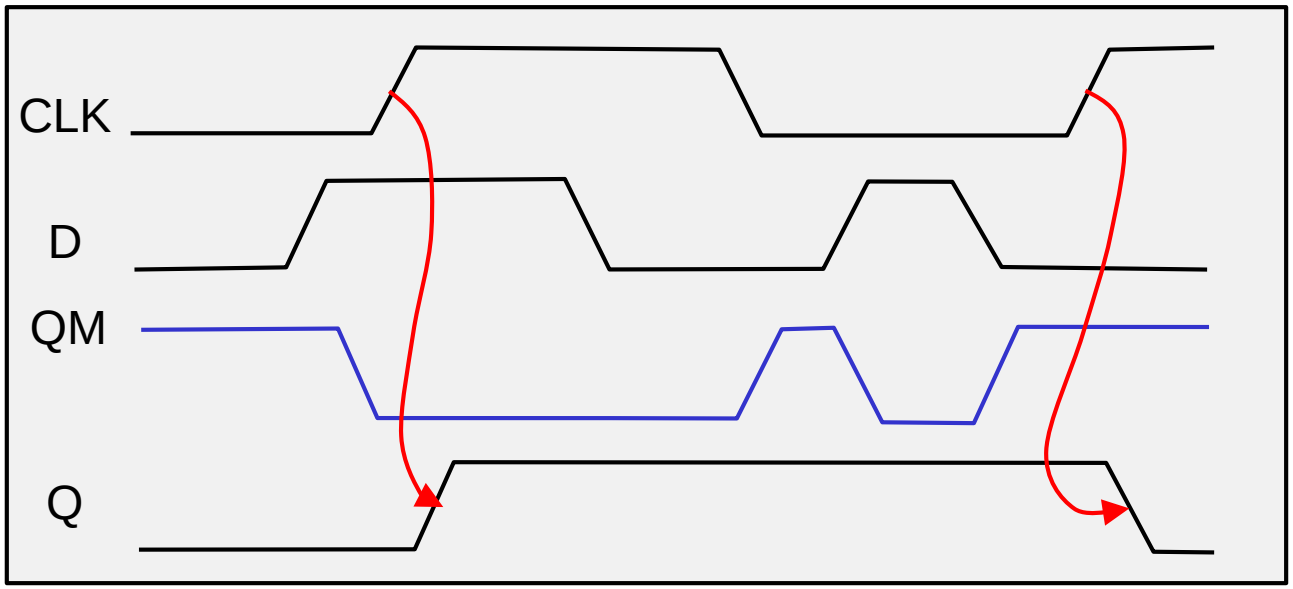
- D Flip-flop is built from two D latches



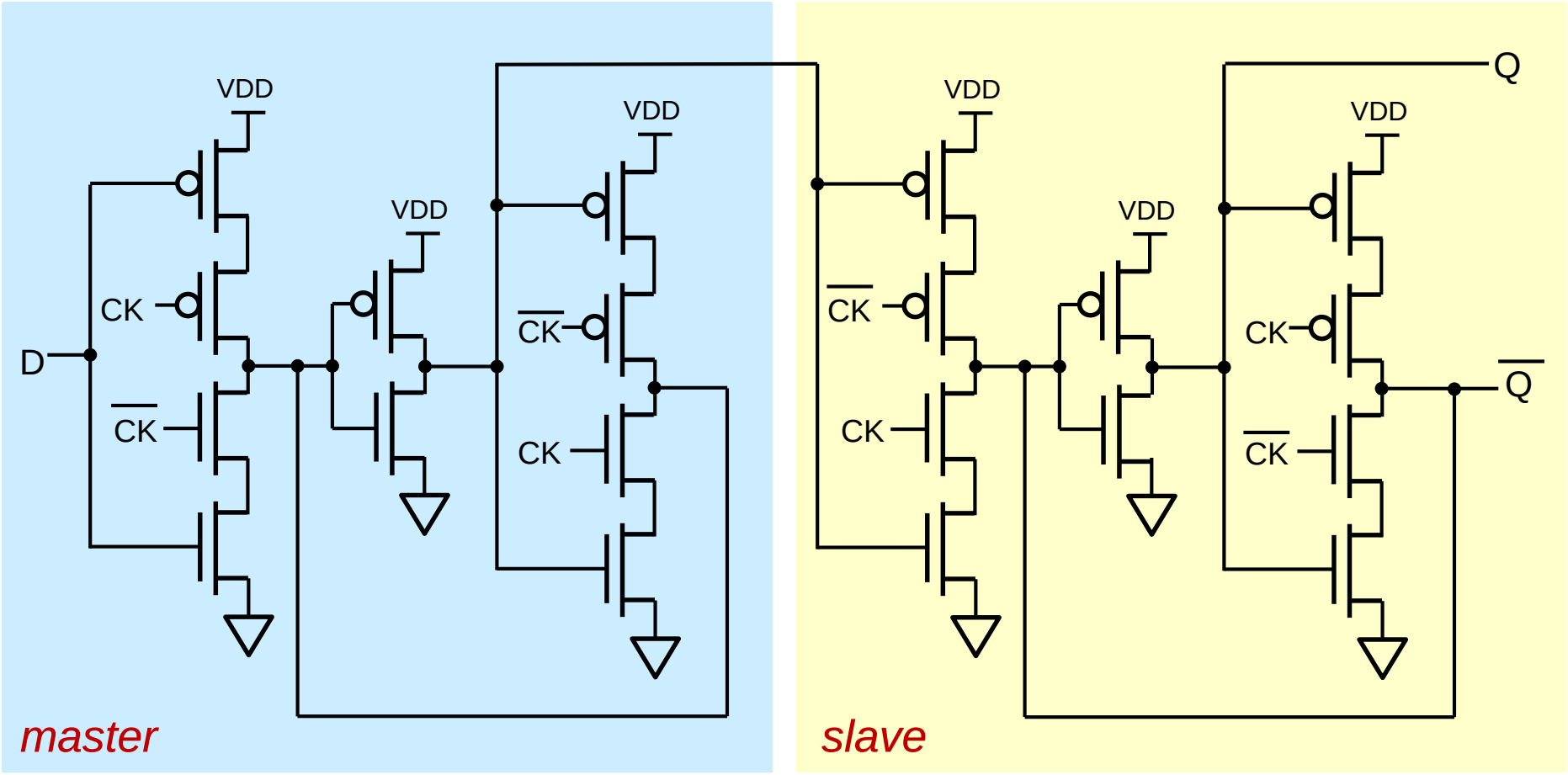


D Flip-flop Operation





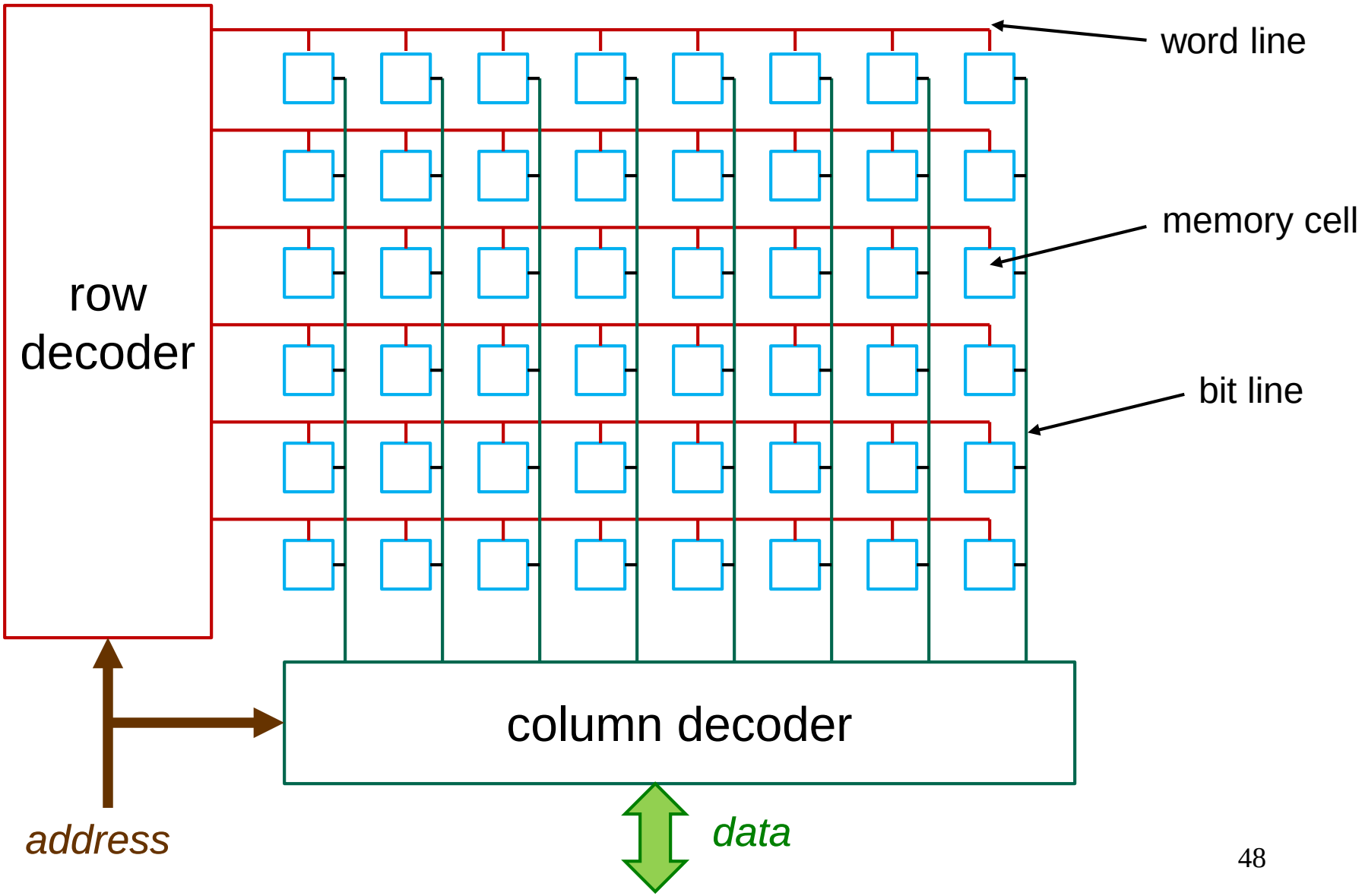
Another D-Flip-flop Implementation



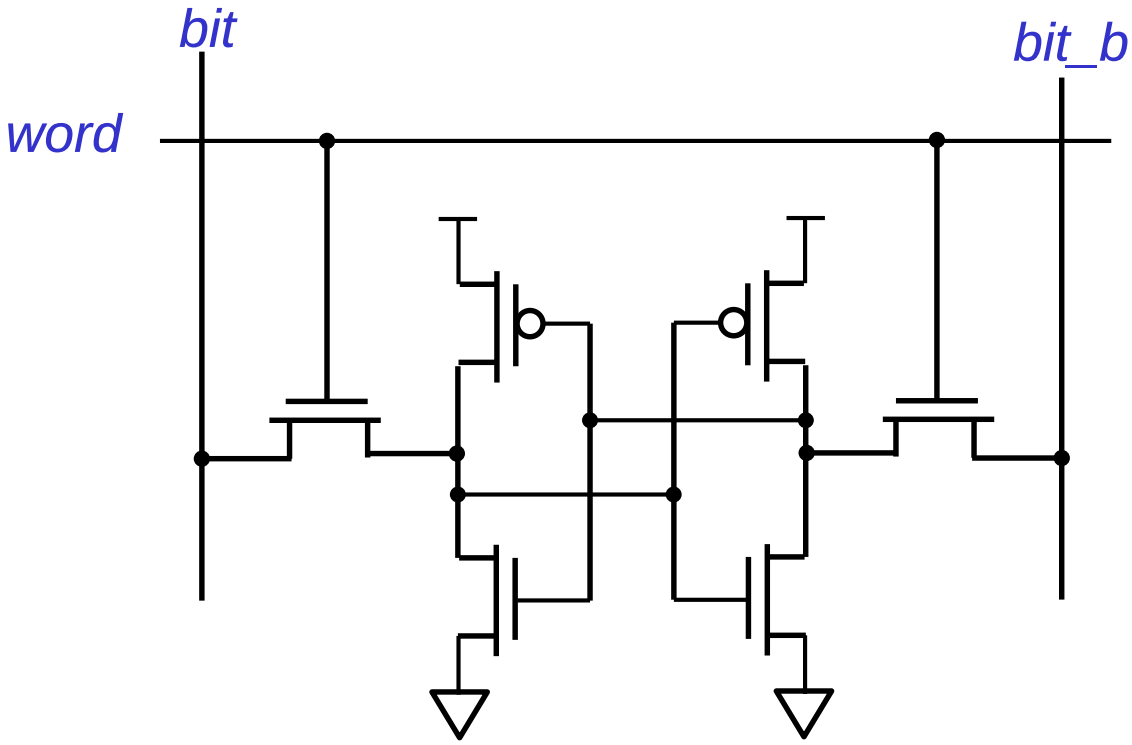
CMOS Memory

- Latches and flip-flops are used to hold temporary values in CMOS data paths and control circuitry
  - register arrays
  - finite state machines
- Not suitable for large memory arrays
  - too much area (D flip-flop uses 20 transistors)
  - too much power
- Memory arrays are designed using
  - smallest possible unit cell
  - regular layout
  - yields very high density and simple design
- CMOS static RAM (SRAM) - 6 transistors per bit
- CMOS dynamic RAM (DRAM) - 1-3 transistors per bit

Memory Array Architecture

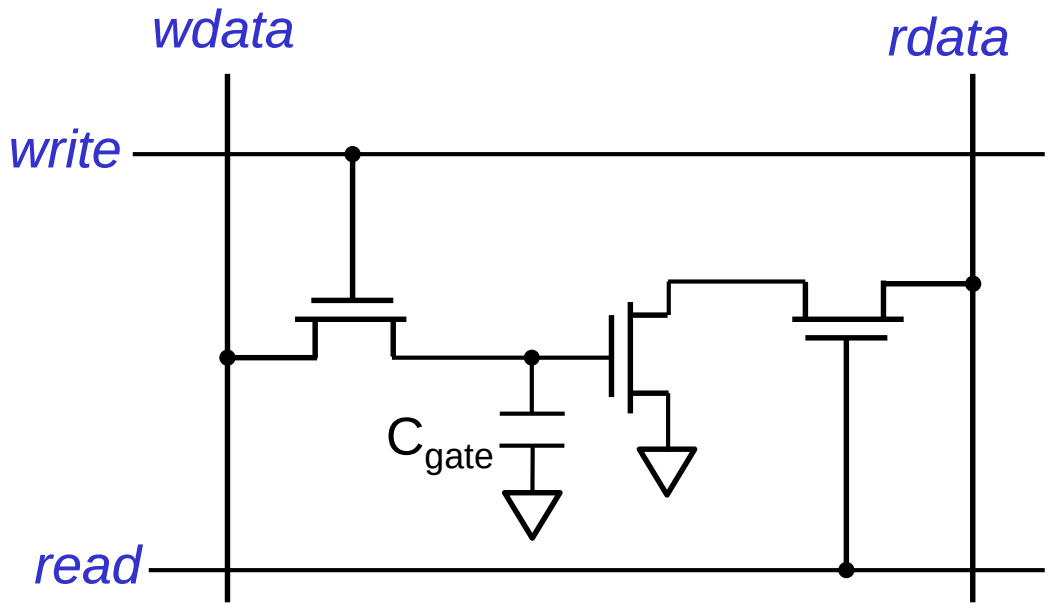


6-T SRAM Cell



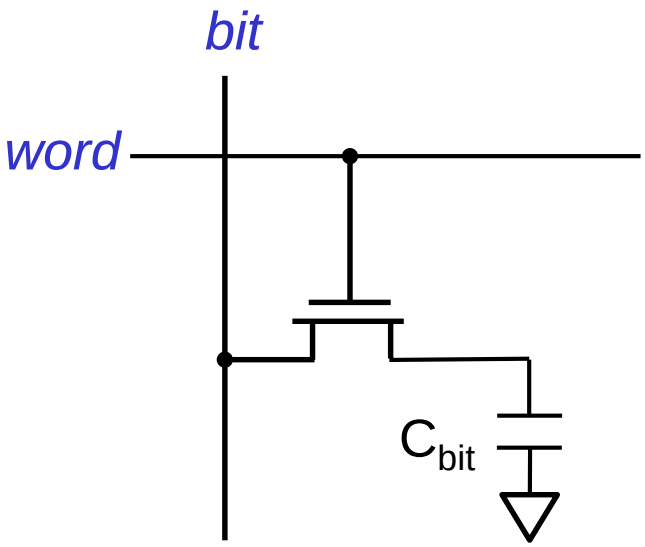
- Used in most commercial chips
  - Data stored in cross-coupled inverters
- Read: precharge *bit*, *bit_b*, then raise *word* line
- Write: drive data onto *bit*, *bit_b*, then raise *word* line

3-T DRAM Cell



- Data stored as charge on gate capacitance C_{gate}
- Read: precharge *rdata*, then raise *read* line
- Write: drive data onto *wdata*, then raise *write* line
- Charge will eventually leak away
  - cell must be periodically refreshed (read followed by re-write)
  - requires more complex memory controller

1-T DRAM Cell



Read: precharge *bit*, then raise *word* line

Write: drive data onto *bit*, then raise *word* line

- Highest density - cell used in commercial DRAM chips
  - Data stored as charge on capacitor C_{bit}
- Charge eventually leaks away – requires refresh
- Read operation is destructive
  - read generates only a small change in voltage of bit line due to charge sharing - requires sense amplifier on bit lines
- Not usually implemented in CMOS
  - requires special process to make a (physically) small capacitor with large capacitance