

ECE 6473
Lecture 6
Date: 10/19/2015

Shaloo Rakheja
Assistant Professor
Electrical and Computer Engineering, NYU

Reading

- Power point and hand-written lecture notes posted on newclasses.nyu.edu
- Sections 7.1, 7.2, 7.3.1 from Digital Integrated Circuits by Jan M. Rabaey et al.

Homework # 3

- **Due on 10/19 without any exceptions. If the server is down, please email your homework to the TAs.**

Homework # 4

- Due on 10/26/2015
- Cadence design problems + dynamic logic

Exam on 11/02

- Everything covered up to 10/26/2015.
- On 10/26/2015, we will go over the project, so that will not be included in the test. We will also discuss doubts and recap the materials. Examples from recap materials could be asked on the midterm.
- 2-page cheat sheet is allowed. You will have to attach the cheat sheet with your answer sheet.
- Class will be split into three sections and exam will be conducted in separate classrooms. Room information will be given on 10/26.

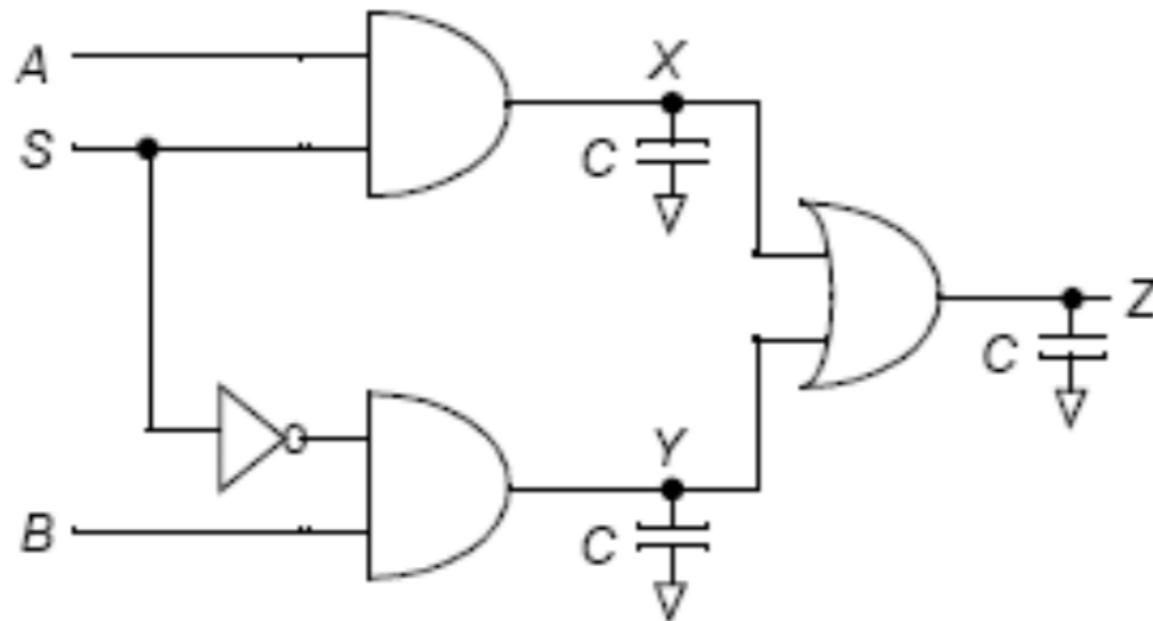
Content

1. Power dissipation

2. Sequential logic

2. Static random access memory (*not on
the mid-term*)

Power dissipation



$$\Pr(X = 1) = p_X = p_A p_S$$

$$\Pr(Y = 1) = p_Y = p_B (1 - p_S)$$

$$\begin{aligned}\Pr(Z = 1) &= \Pr(X = 1 \text{ or } Y = 1) = p_X + p_Y - \Pr(X = 1 \text{ and } Y = 1) \\ &= p_X + p_Y - \underbrace{\Pr(X = 1 | Y = 1)}_0 p_Y = p_X + p_Y = p_A p_S + p_B (1 - p_S)\end{aligned}$$

If the signal probabilities are 0.5 =>

$$p_z = 0.5 \times 0.5 + 0.5 \times (1 - 0.5) = 0.5$$

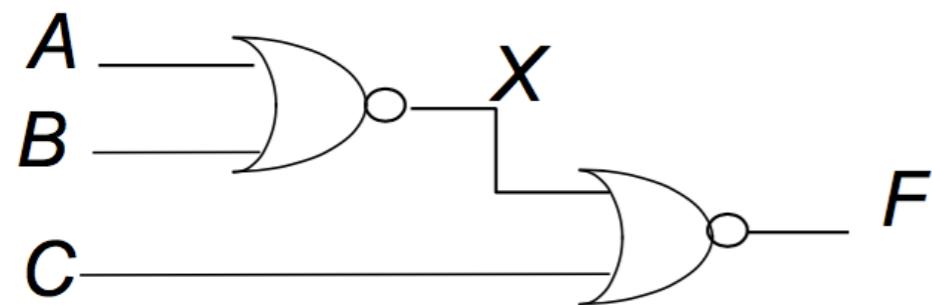
Result verification

$$X = A.S, Y = B.\bar{S}, Z = A.S + B.\bar{S}$$

A	B	S	F	X	Y
0	0	0	0	0	0
0	0	1	0	0	0
0	1	0	1	0	1
0	1	1	0	0	0
1	0	0	0	0	0
1	0	1	1	1	0
1	1	0	1	0	1
1	1	1	1	1	0

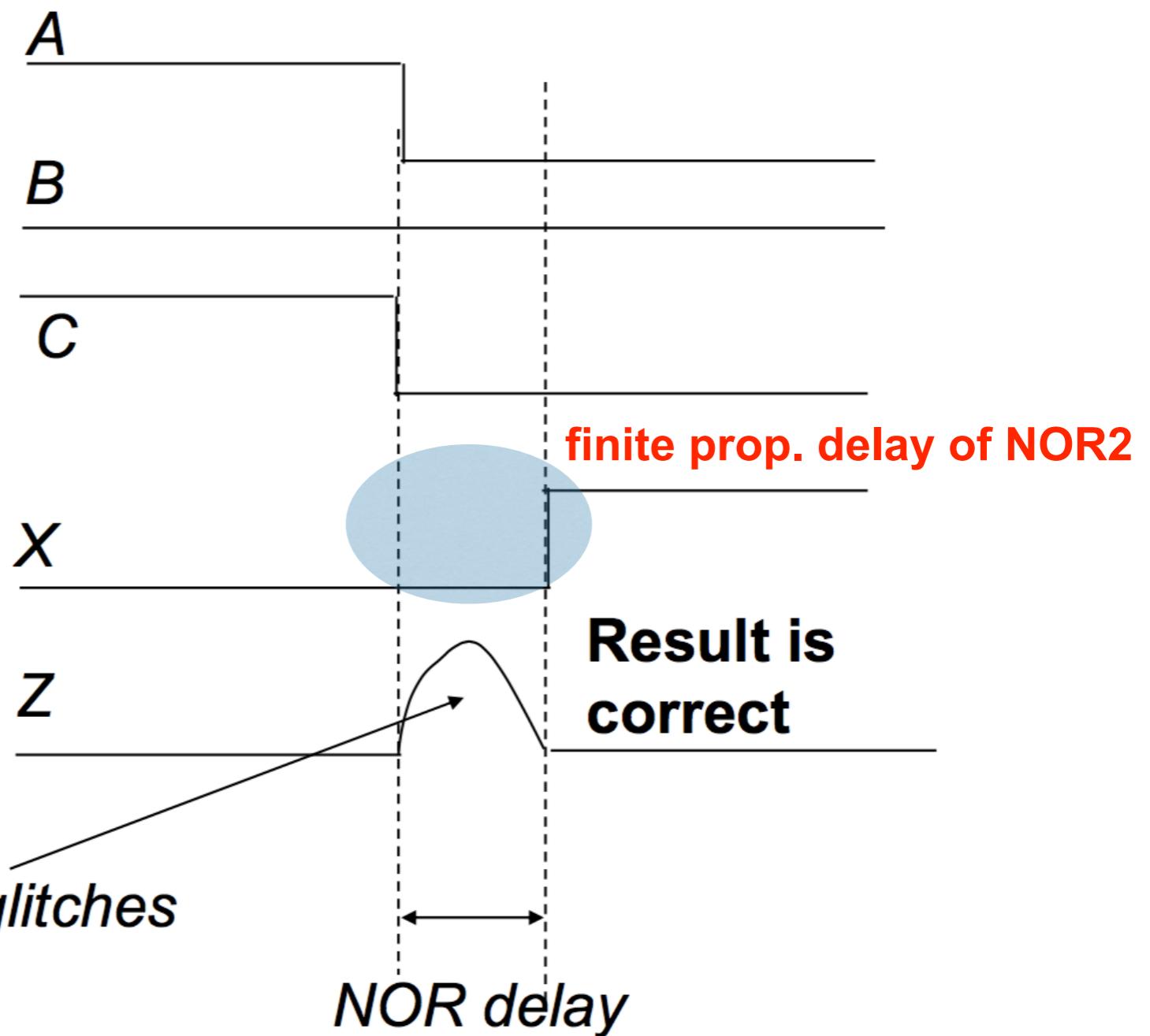
$$p_Z = 4/8 = 0.5$$

Glitches or dynamic hazards

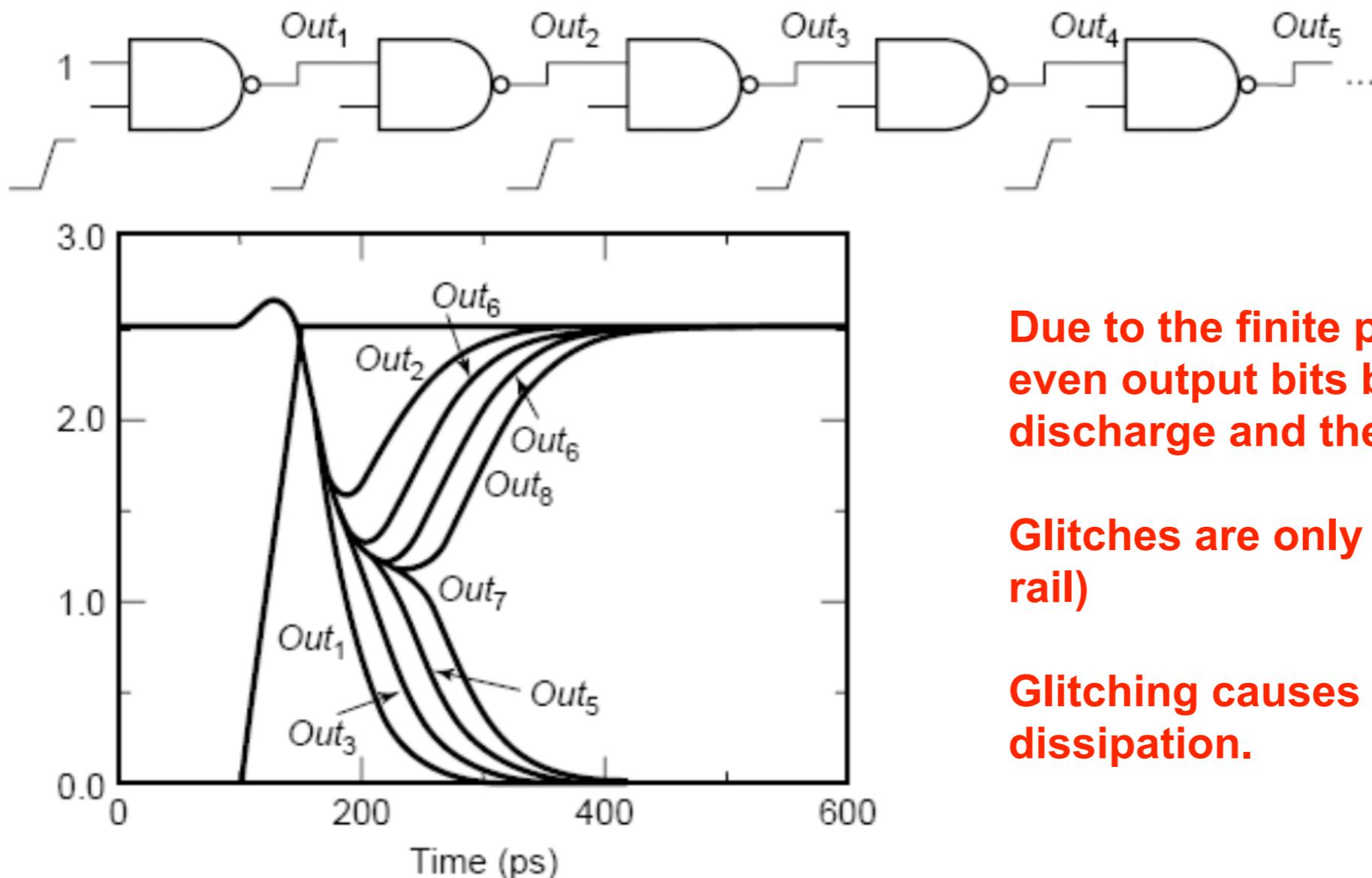


**Dynamic Hazards or
glitch result in additional
power dissipation**

Dynamic Hazards or glitches



Glitches or dynamic hazards



Due to the finite prop. delay, the even output bits begin to discharge and the voltage drops.

Glitches are only partial (not rail to rail)

Glitching causes extra power dissipation.

A chain structure tends to have higher glitching power

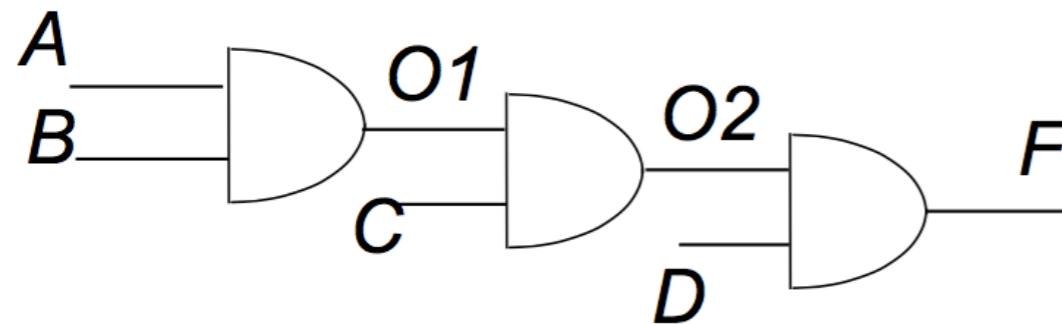
Techniques to reduce switching activity

- ❑ Logic restructuring
- ❑ Input ordering
- ❑ Glitch reduction by path balancing

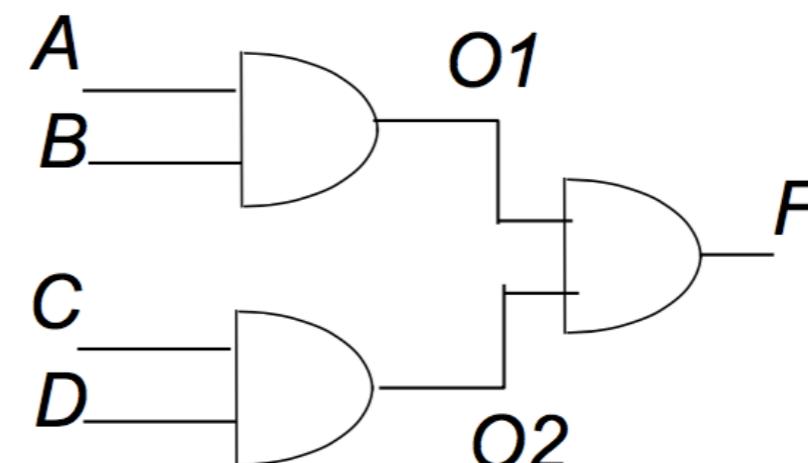
In addition to these techniques, physical capacitance of the circuit must be reduced to reduce the dynamic power dissipation.

Physical capacitance depends on (i) circuit style, (ii) transistor sizing, (iii) placement and routing, (iv) architectural optimizations.

Logic restructuring



Chain structure



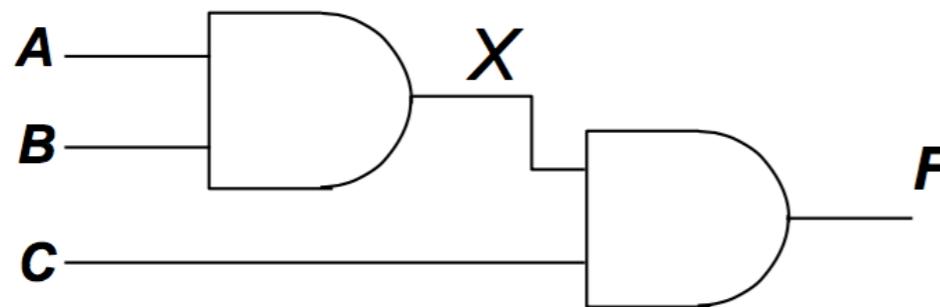
Tree structure

	Chain Structure			Tree Structure		
	O1	O2	F	O1	O2	F
p_1	1/4	1/8	1/16	1/4	1/4	1/16
p_0	3/4	7/8	15/16	3/4	3/4	15/16
$\alpha_{0 \rightarrow 1}$	3/16	7/64	15/256	3/16	3/16	15/256

Chain has lower activity but higher glitching

Input reordering

$$p_A = 0.5, p_B = 0.2, p_C = 0.1$$



At Node X:

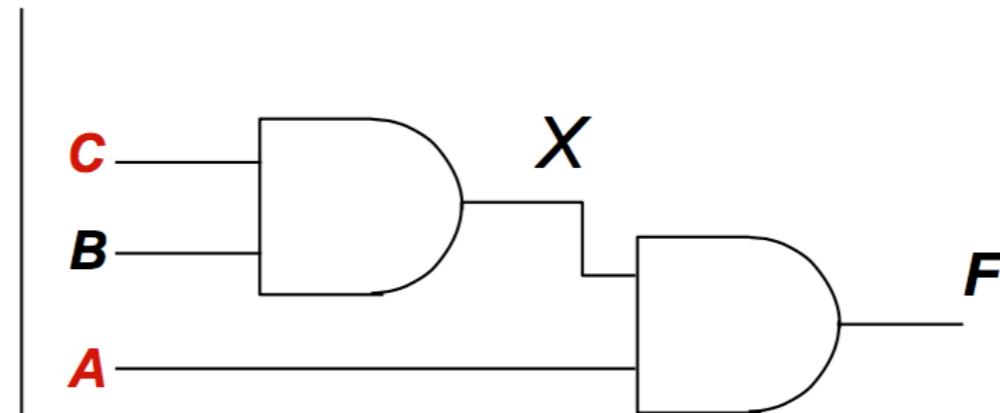
$$p_X = p_A p_B = 0.5 \times 0.2 = 0.1$$

$$\alpha_{0 \rightarrow 1} = (1 - p_X) p_X = 0.1 \times 0.9 = 0.09$$

At Node F:

$$p_F = p_X p_C = 0.1 \times 0.1 = 0.01$$

$$\alpha_{0 \rightarrow 1} = (1 - p_F) p_F = 0.99 \times 0.01 = 0.0099$$



At Node X:

$$p_X = p_C p_B = 0.1 \times 0.2 = 0.02$$

$$\alpha_{0 \rightarrow 1} = (1 - p_X) p_X = 0.02 \times 0.98 = 0.0196$$

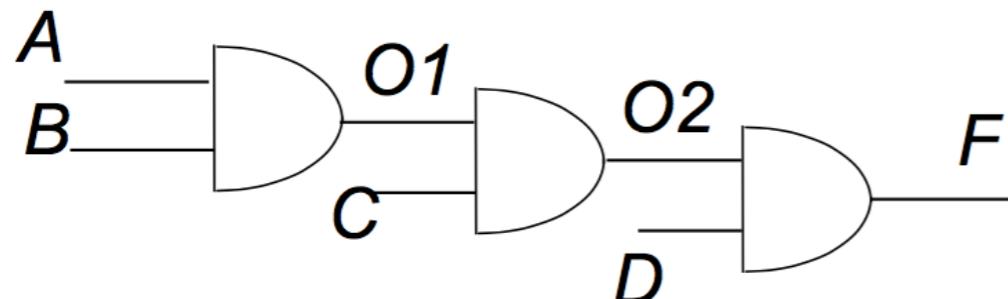
At Node F:

$$p_F = p_X p_Z = 0.02 \times 0.5 = 0.01$$

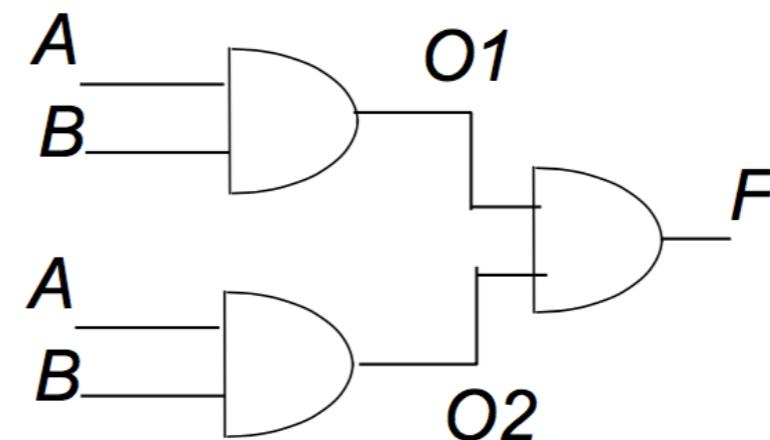
$$\alpha_{0 \rightarrow 1} = (1 - p_F) p_F = 0.99 \times 0.01 = 0.0099$$

It may be beneficial to introduce signals with higher transition probability later at the logic chain

Path balancing for glitches



Chain structure



Tree structure

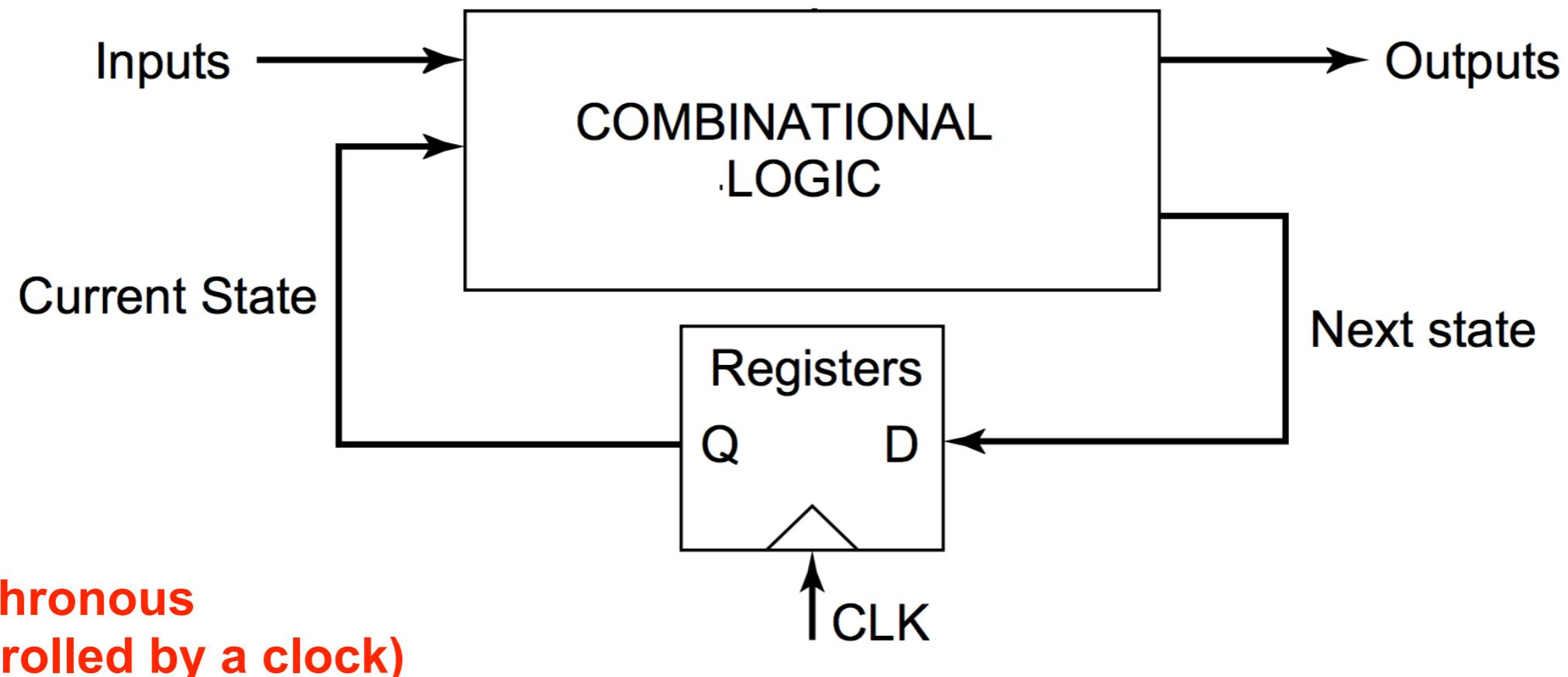
Balancing the path ensures that the signal arrival time at the inputs are closer to each other.

This is very useful particularly when the individual logic elements have long delay

Equalize the path lengths through logic

Sequential logic

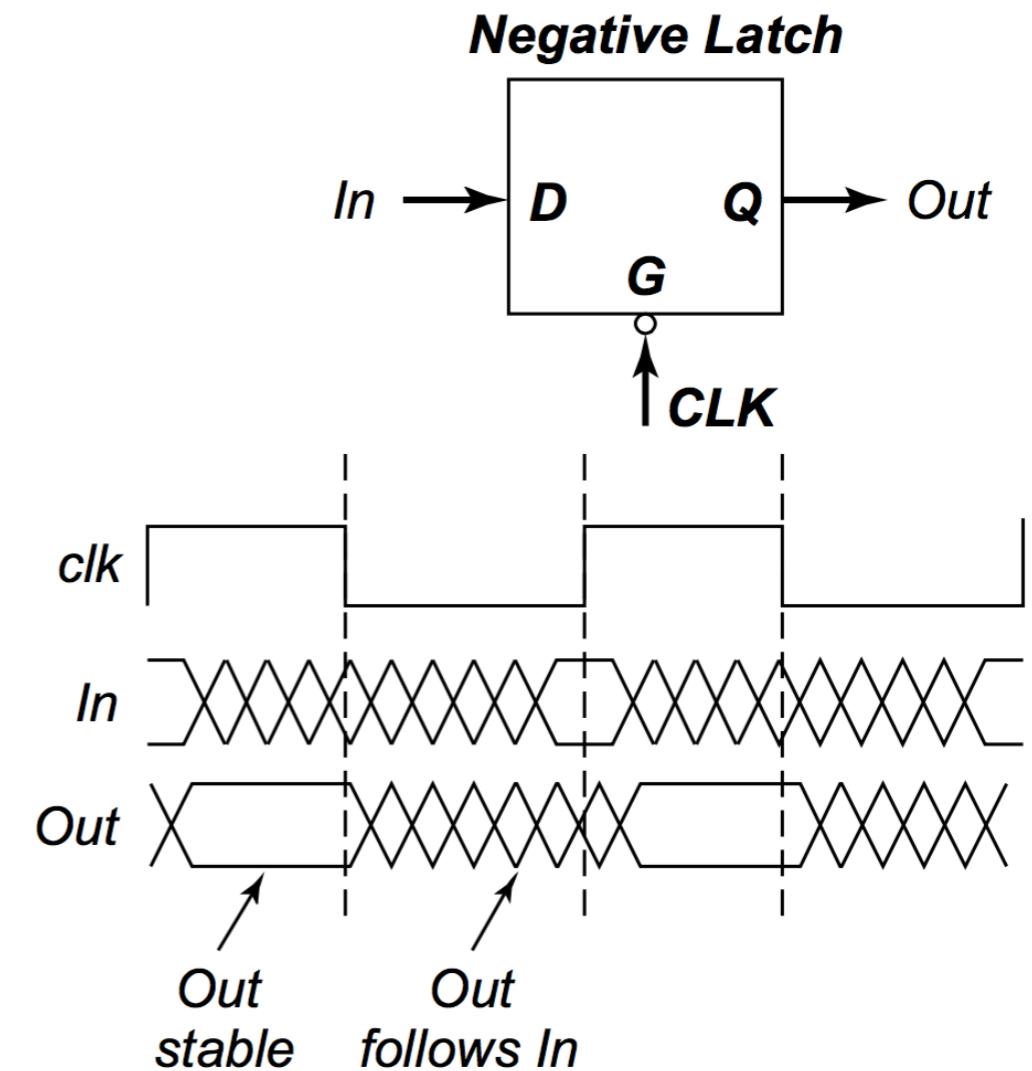
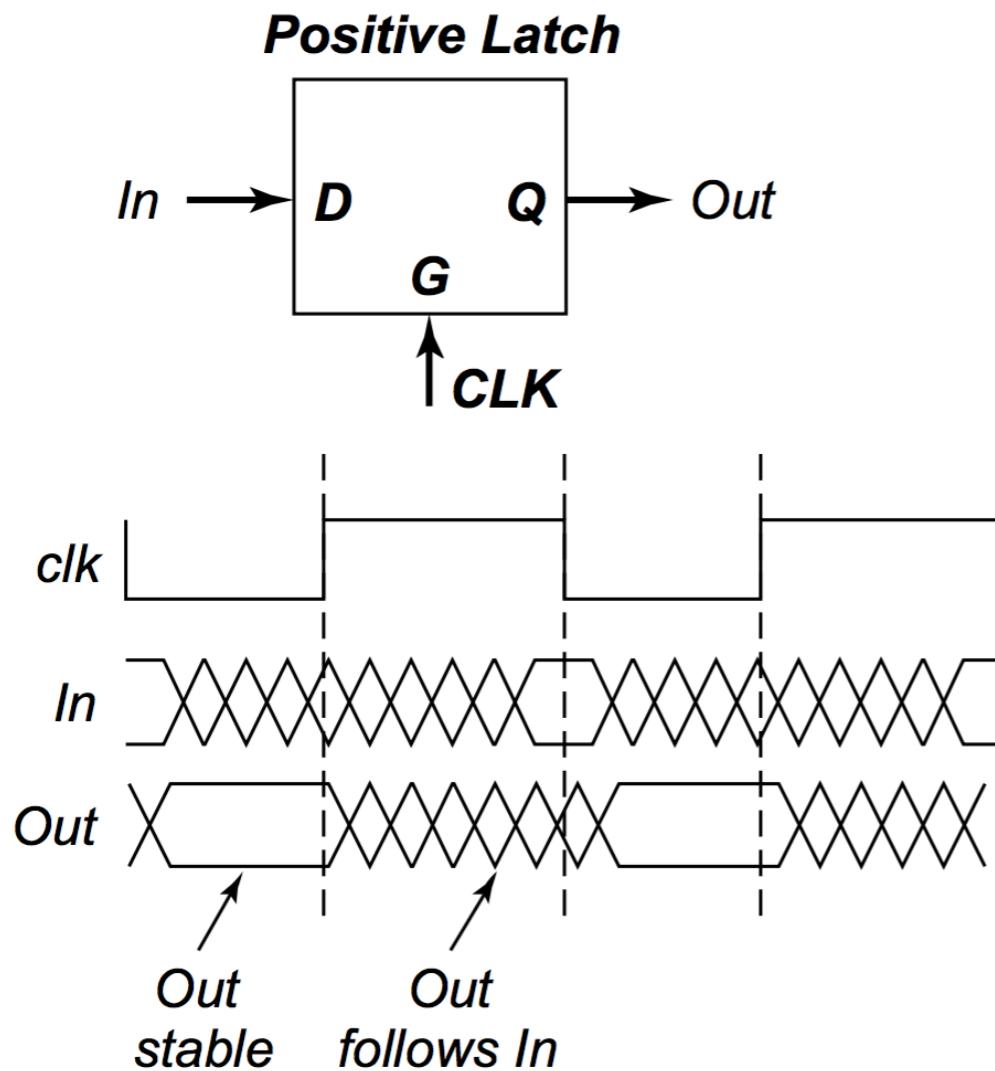
Sequential Logic



2 storage mechanisms

- positive feedback
- charge-based

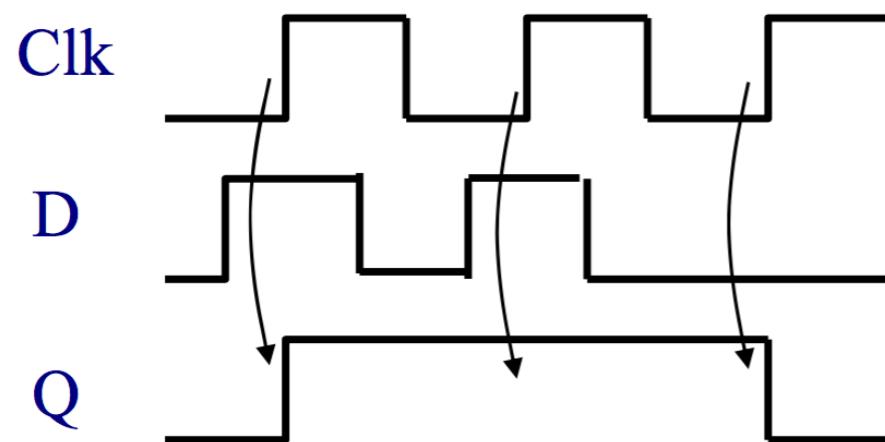
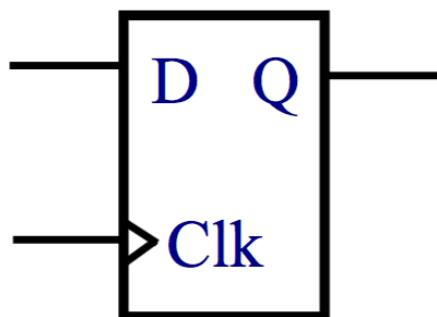
Latches: level sensitive



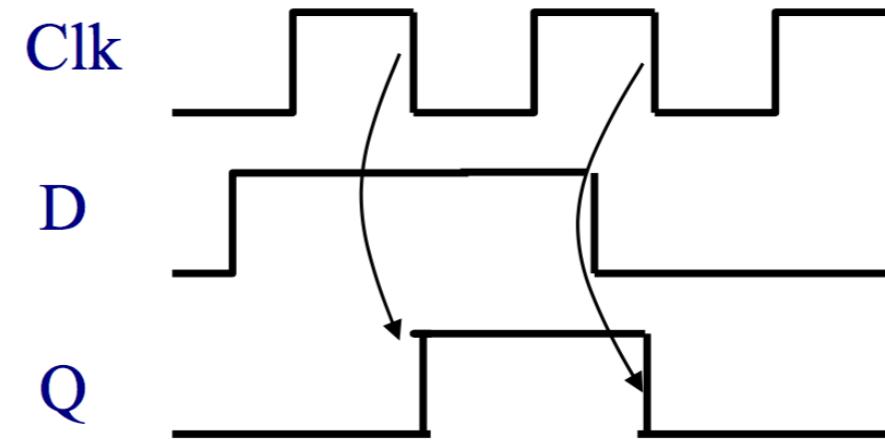
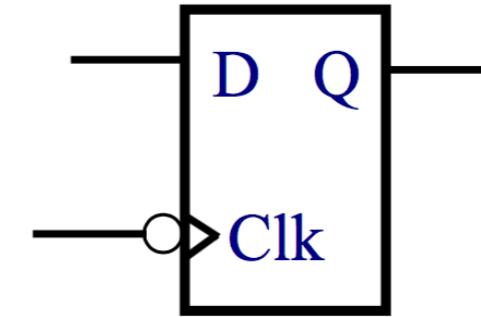
Latch passes D input to Q at high or low level of the clock

Registers/Flip flops: edge sensitive

Positive Edge -Triggered

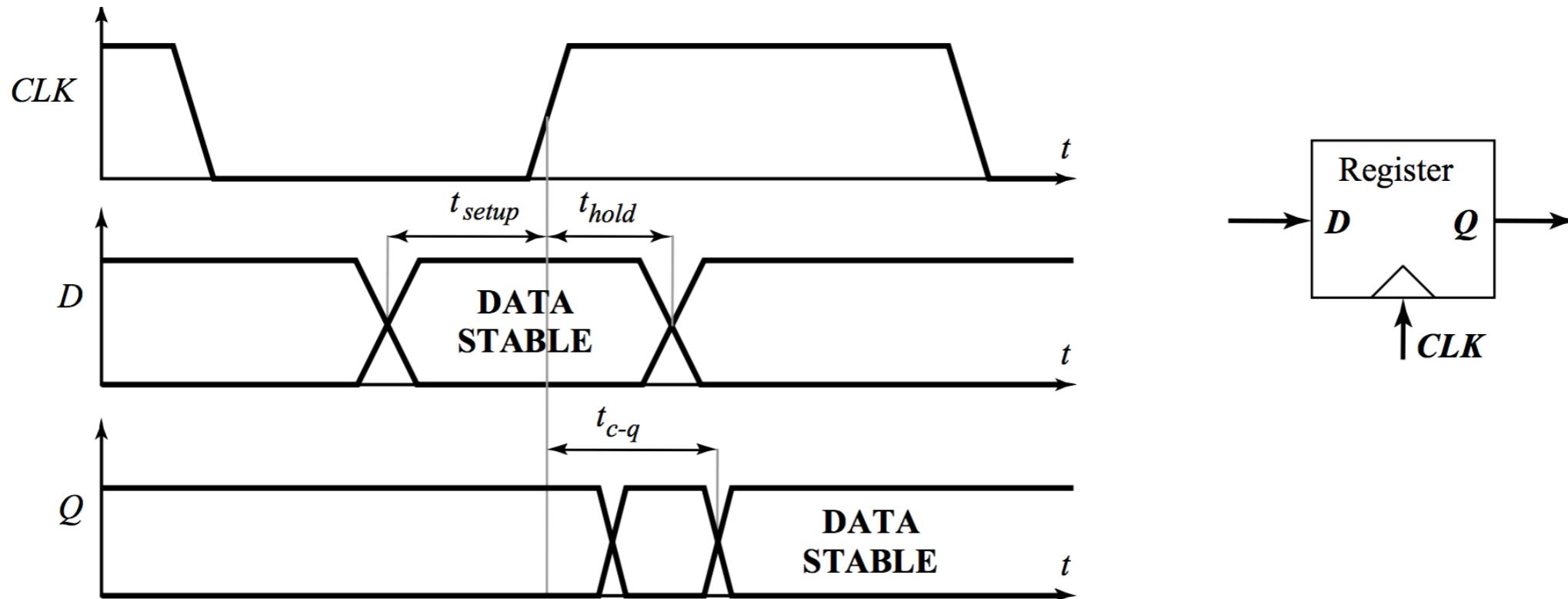


Negative Edge -Triggered



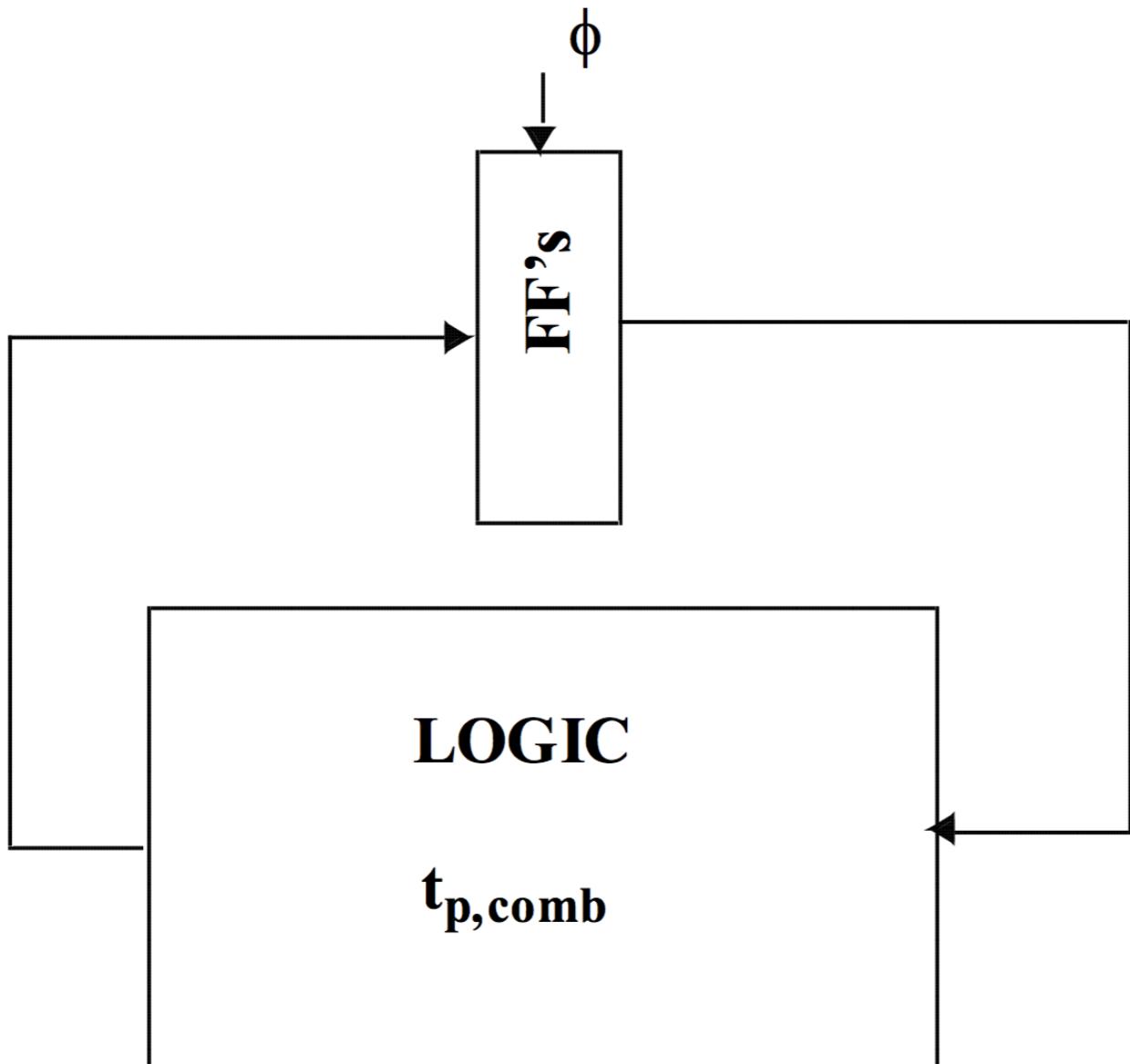
- Register or Flip-flops
D is passed to Q only at the clock edge

Timing definitions



- **Setup time (t_{setup}):** The time that the data input must be valid before CLK arrives
- **Hold time (t_{hold}):** The time the data input must remain valid after the CLK edge
- **Clk-Q-delay (t_{c-q}):** The propagation delay between CLK arrival and the data D being copied to Q

Maximum clock frequency



$$t_{c-q} + t_{p,comb} + t_{setup} = T$$

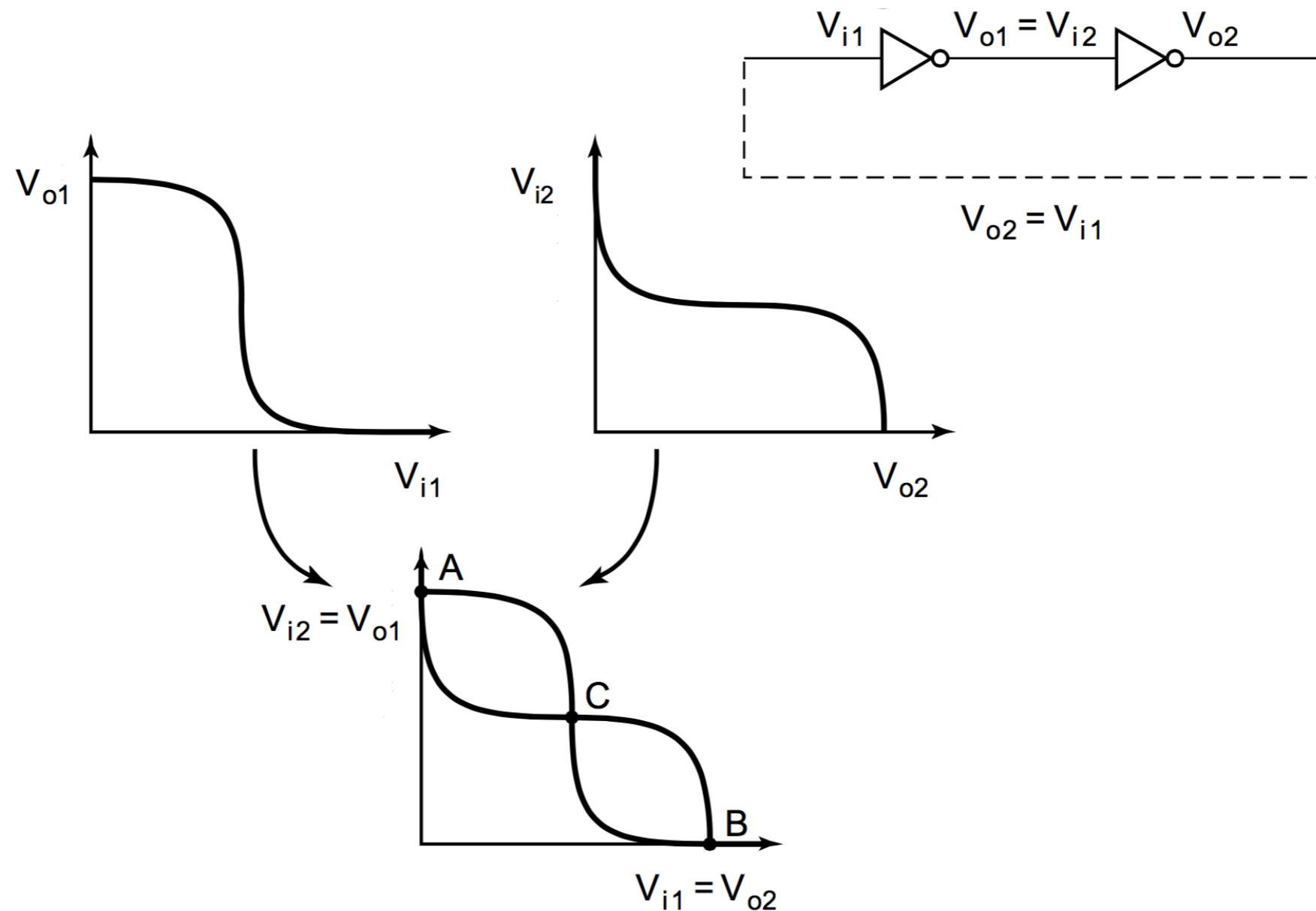
Also:

$$t_{cdreg} + t_{cdlogic} > t_{hold}$$

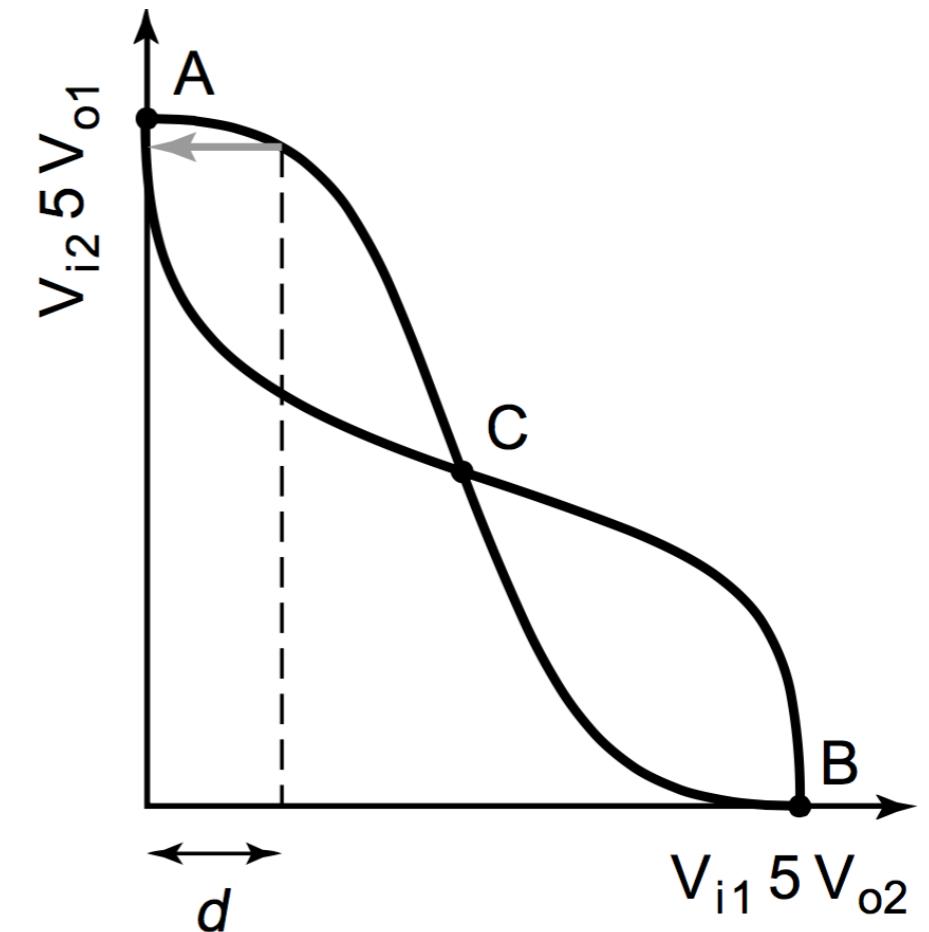
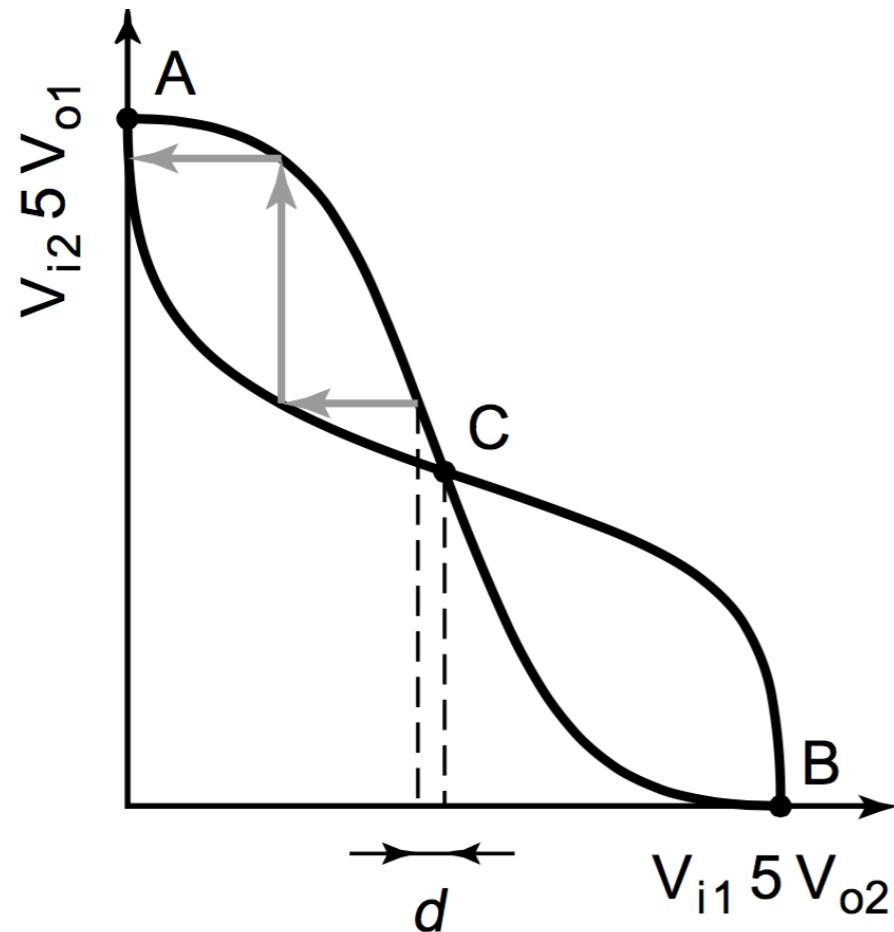
t_{cd} : contamination
delay = minimum
delay

Basic principle of sequential element:

Positive feedback → bistability



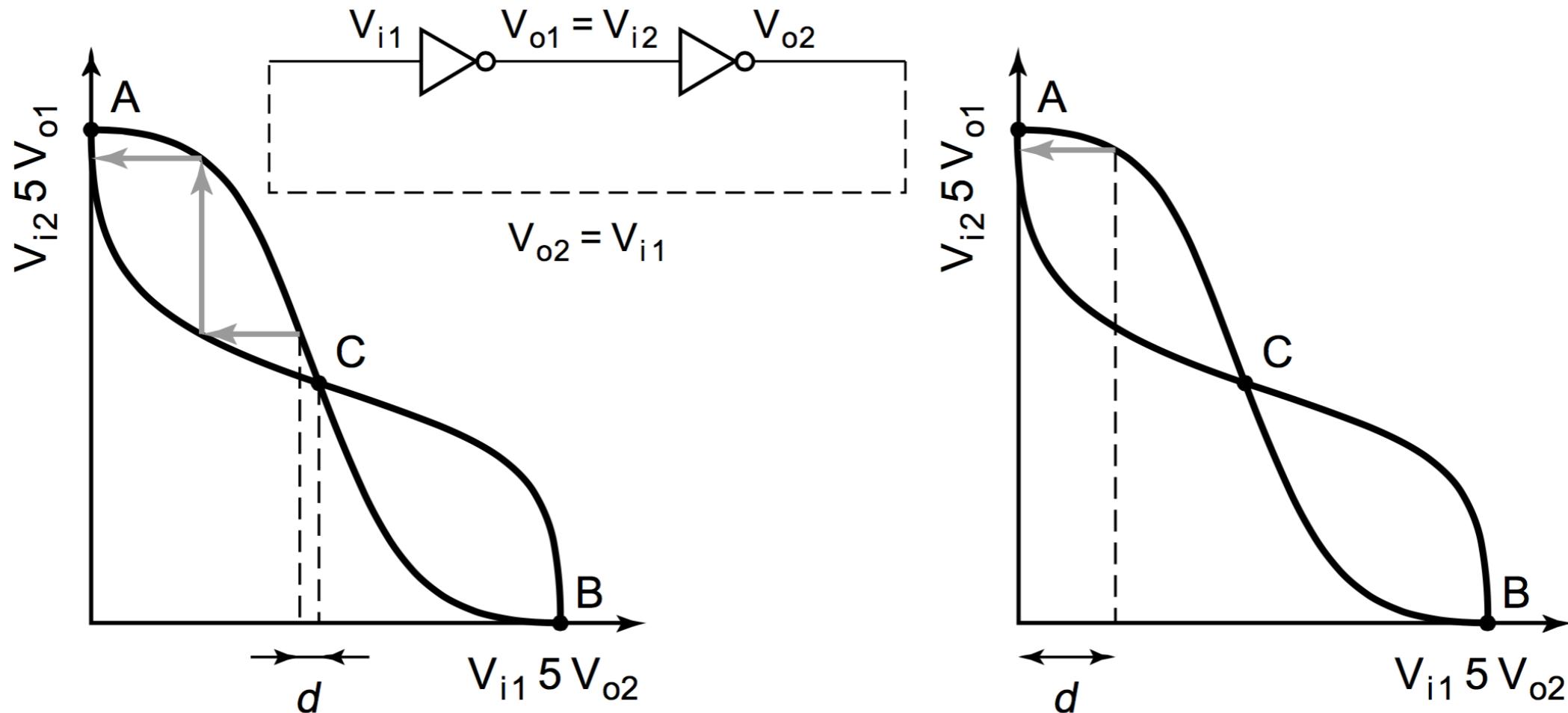
Metastability



Loop Gain should be larger than 1 in the transition region
Loop Gain should be less than 1 in the stable region

A and B are stable points and C is the metastability point

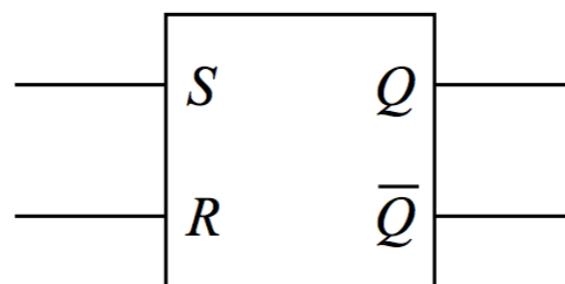
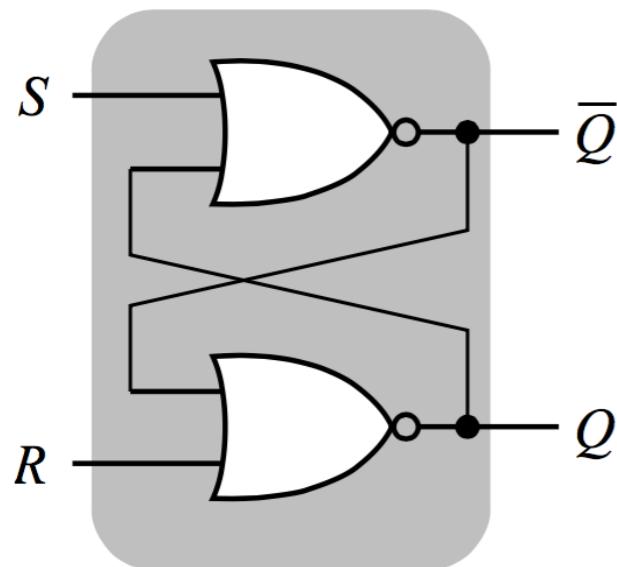
How to change state?



- Keeping the feedback loop and force one input to a region where loop gain is larger than '1'
- Open the feedback loop force the input be in the transition region for the corresponding inverter

Overpowering the feedback loop: NOR set-reset

NOR-based set-reset



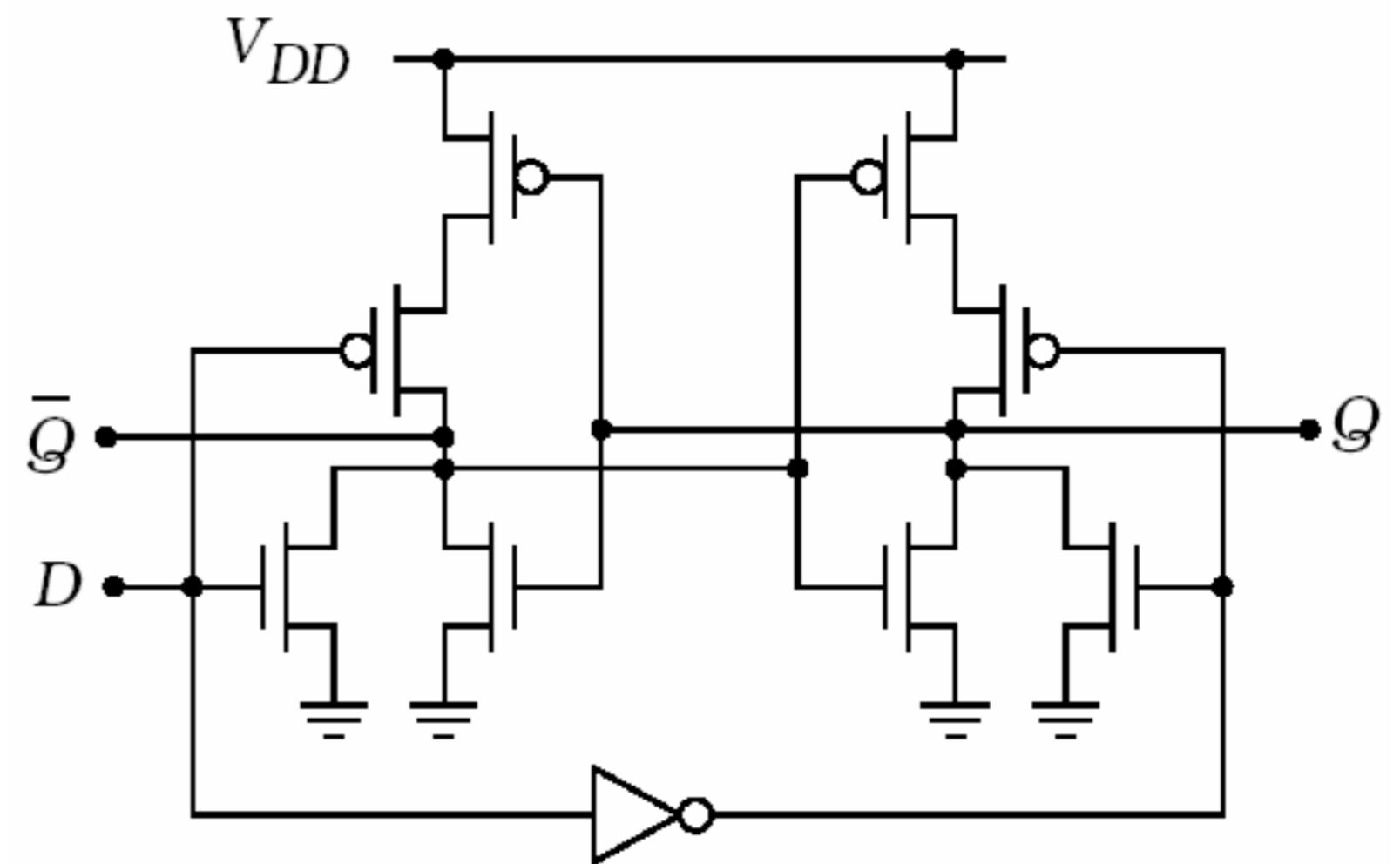
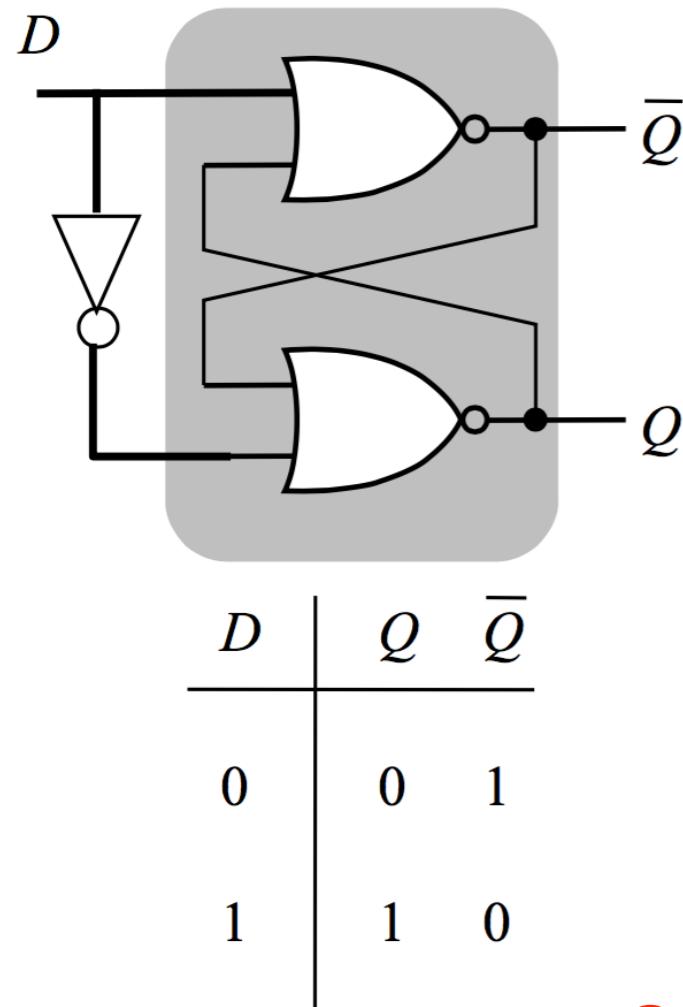
S	R	Q	\bar{Q}
0	0	Q	\bar{Q}
1	0	1	0
0	1	0	1
1	1	0	0

Forbidden State

If both $S = R = 1$ and then they go to 0, the final output becomes unpredictable and depends on whichever input is the last to transition to 0.

Overpowering the feedback loop: NOR-based D latch

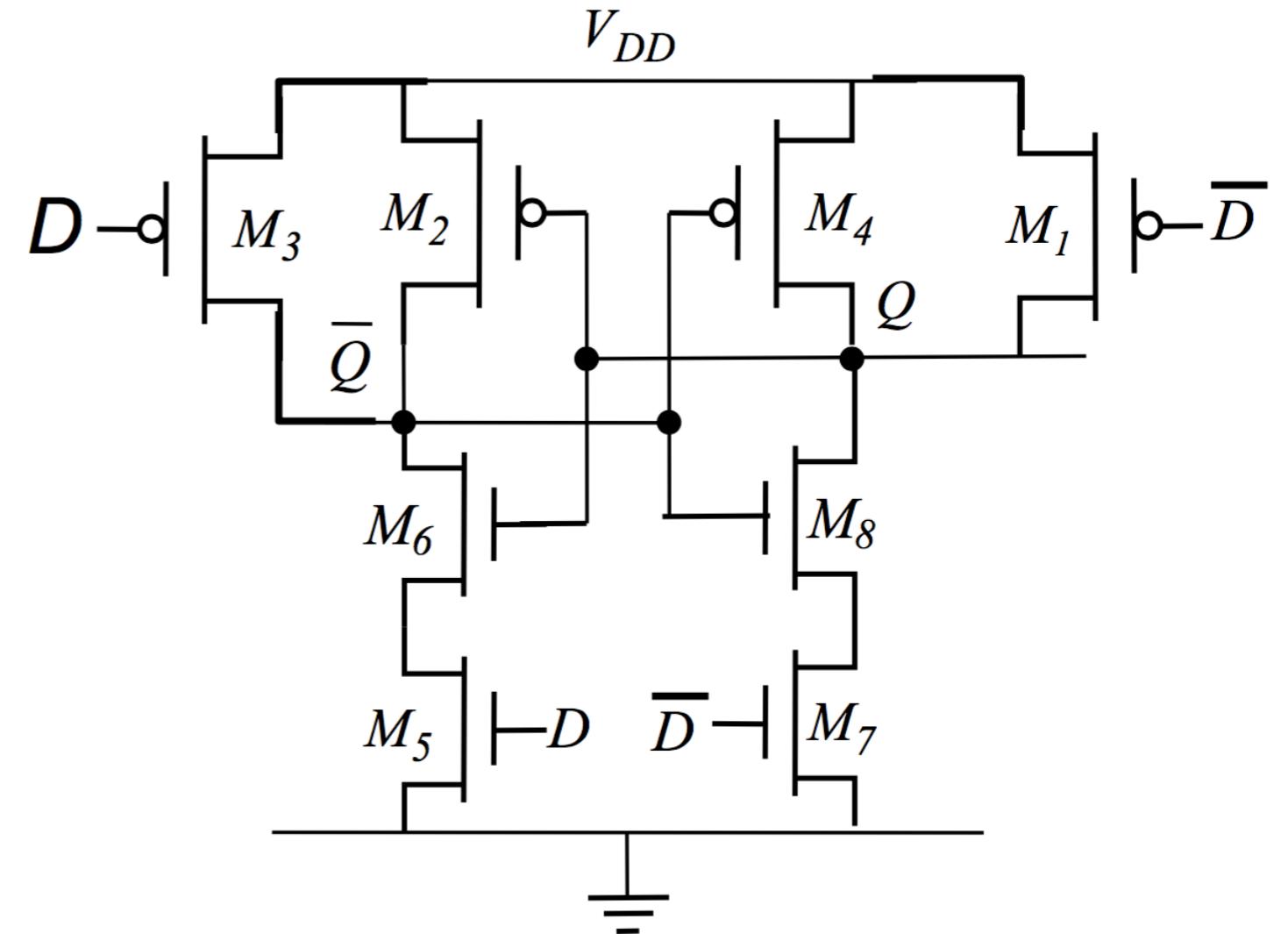
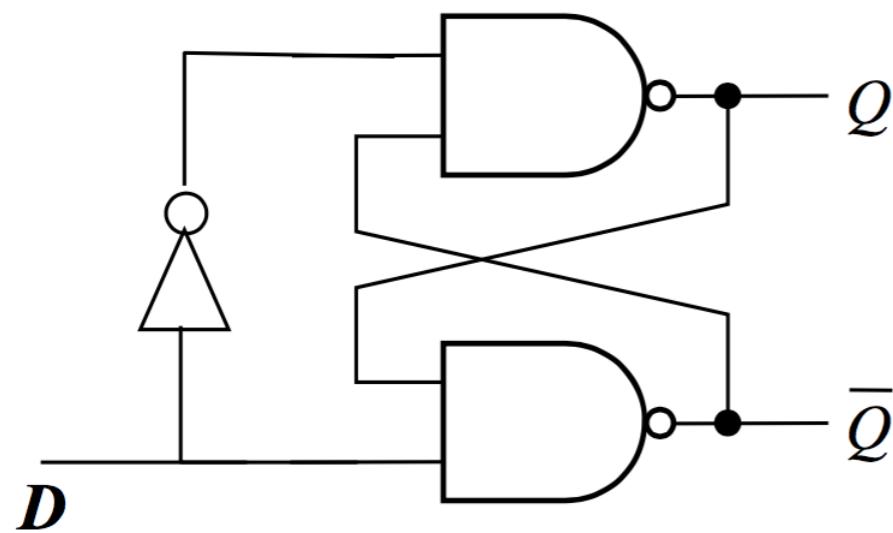
NOR-based D Latch



Only one “Data” signal and its complement are used.
However, this is asynchronous circuit. No clock.

Overpowering the feedback loop: NAND-based D latch

Cross-coupled NANDs



Only one “Data” signal and its complement are used.
However, this is asynchronous circuit. No clock.

D-latch with clock or enable

If CLK or $EN = 0$

\Rightarrow Holds current state

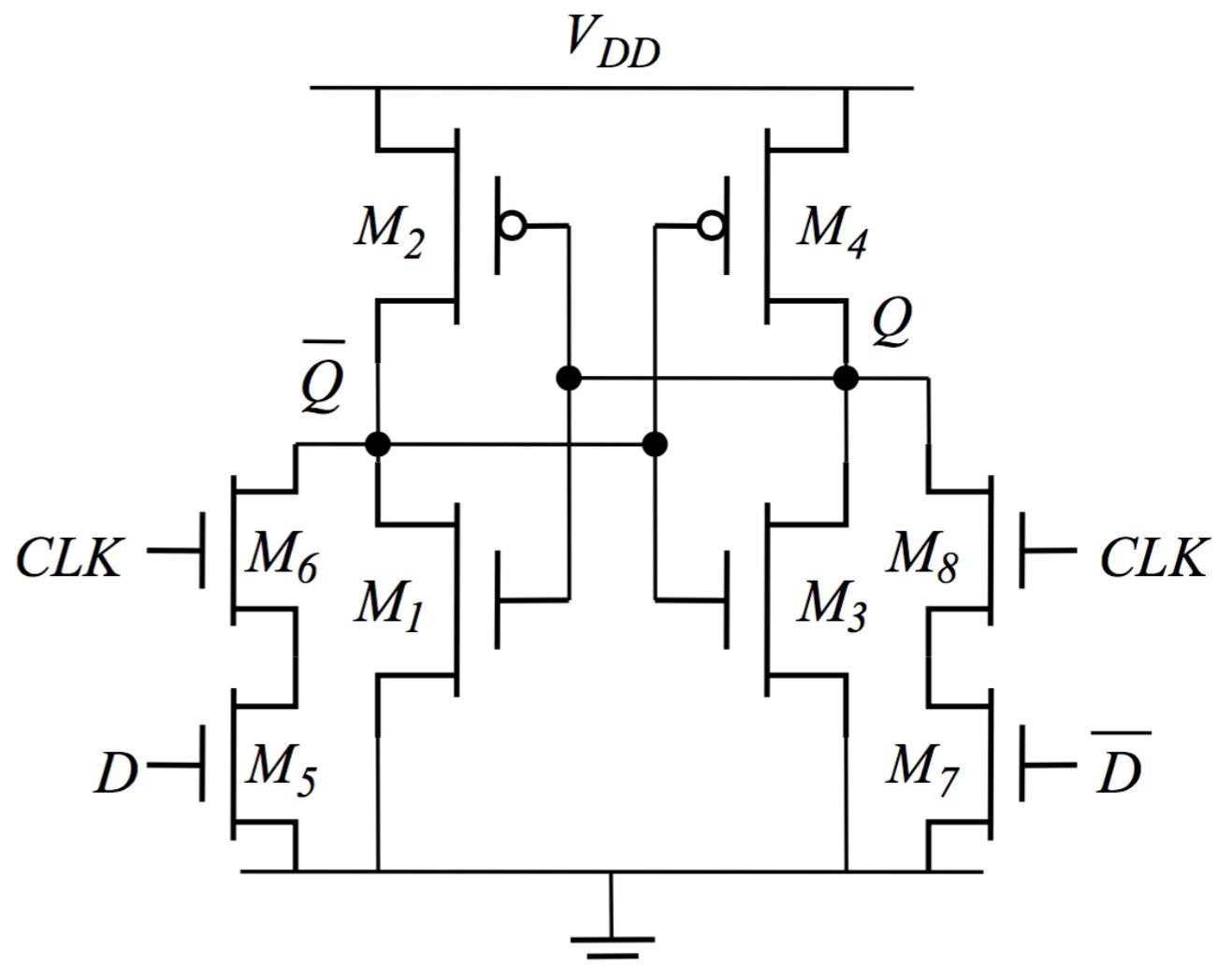
$If CLK$ or $EN = 1 \Rightarrow Q = D$

Add a clock signal for synchronous operation

This is a cross-coupled inverter pair + four extra transistors (2 for clock and 2 for data and its complement)

Clock has an activity factor of ONE. So clock load must be reduced for reducing power dissipation.

Added clock



D-latch with clock or enable

If CLK or $EN = 0$

\Rightarrow Holds current state

If CLK or $EN = 1 \Rightarrow Q = D$

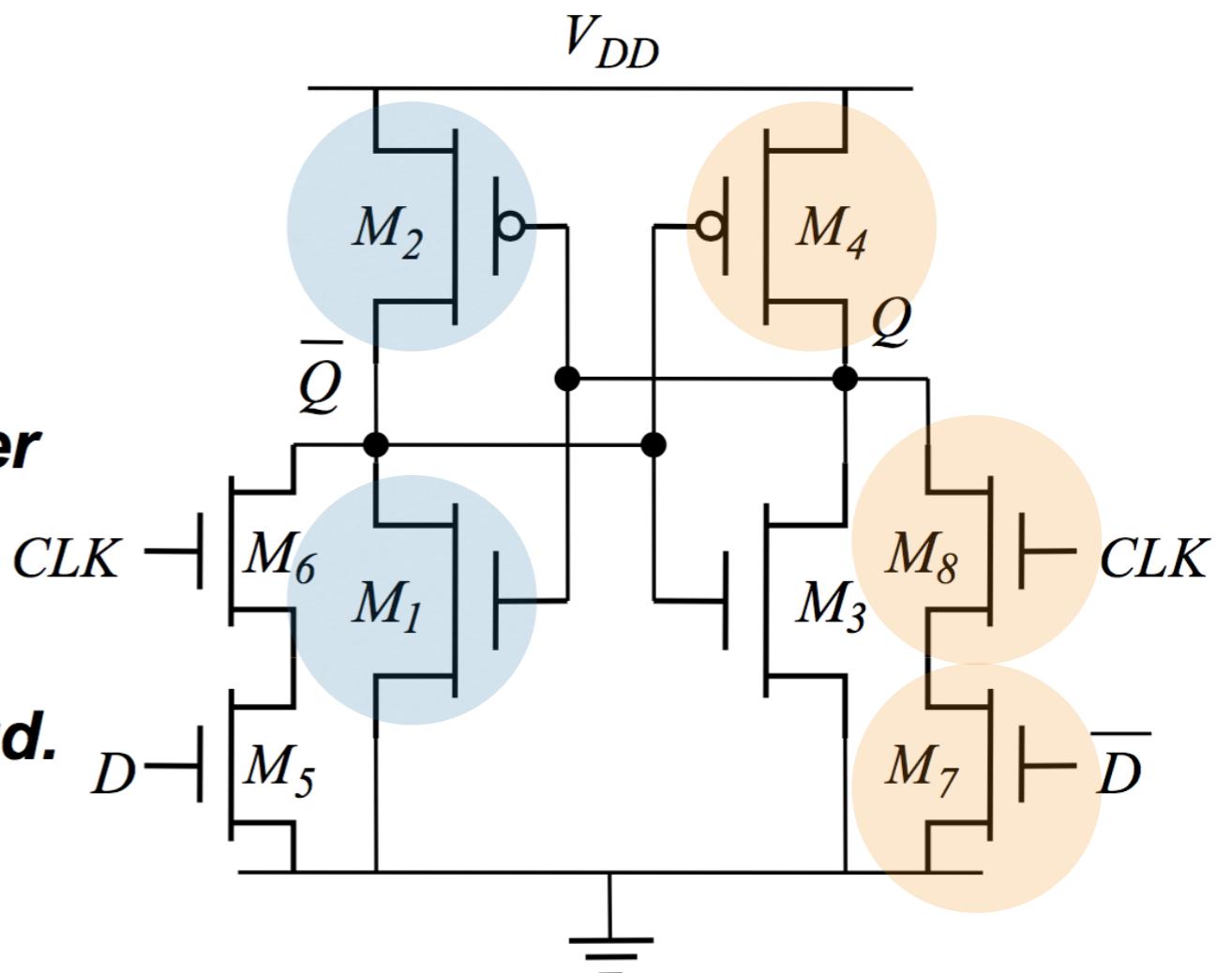
$$CLK = '1' \text{ and } Q = 1, D = 1 \rightarrow 0$$

$M_4 - M_8/M_7$ pseudo-NMOS Inverter

Reduce Q to the switching threshold of $M_2 - M_1$

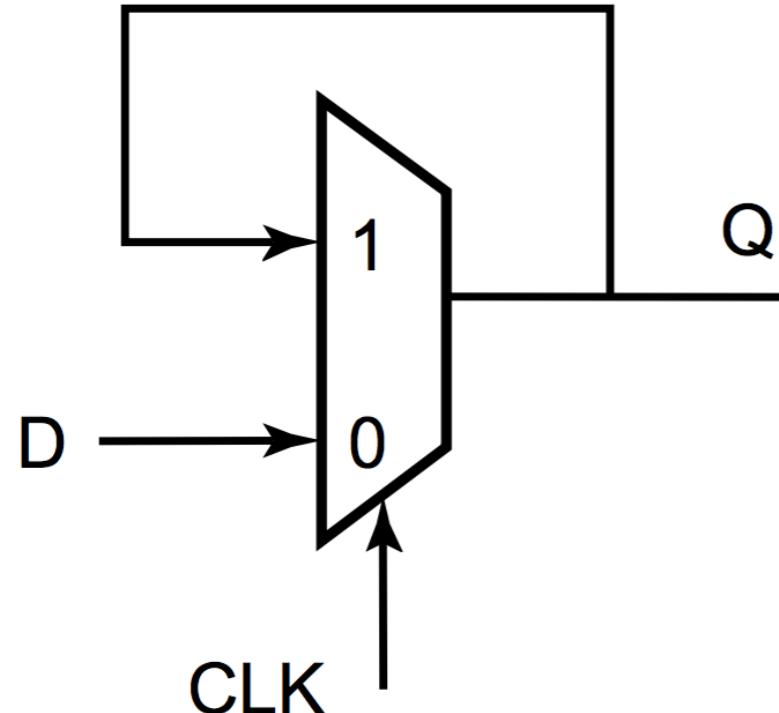
Proper sizing of M_4 - M_8/M_7 is reqd.

Added clock

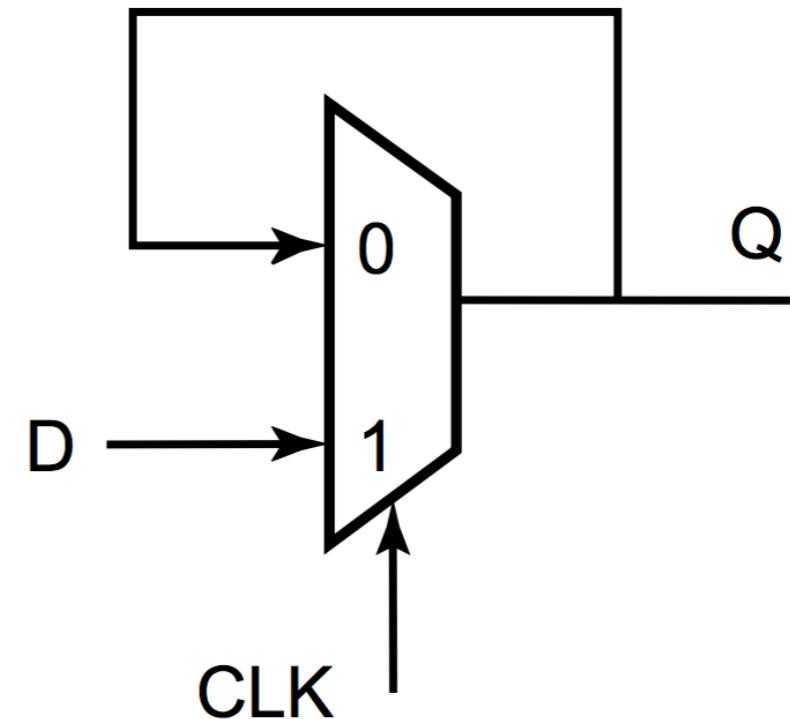


MUX-based latch

Negative latch
(transparent when CLK= 0)



Positive latch
(transparent when CLK= 1)

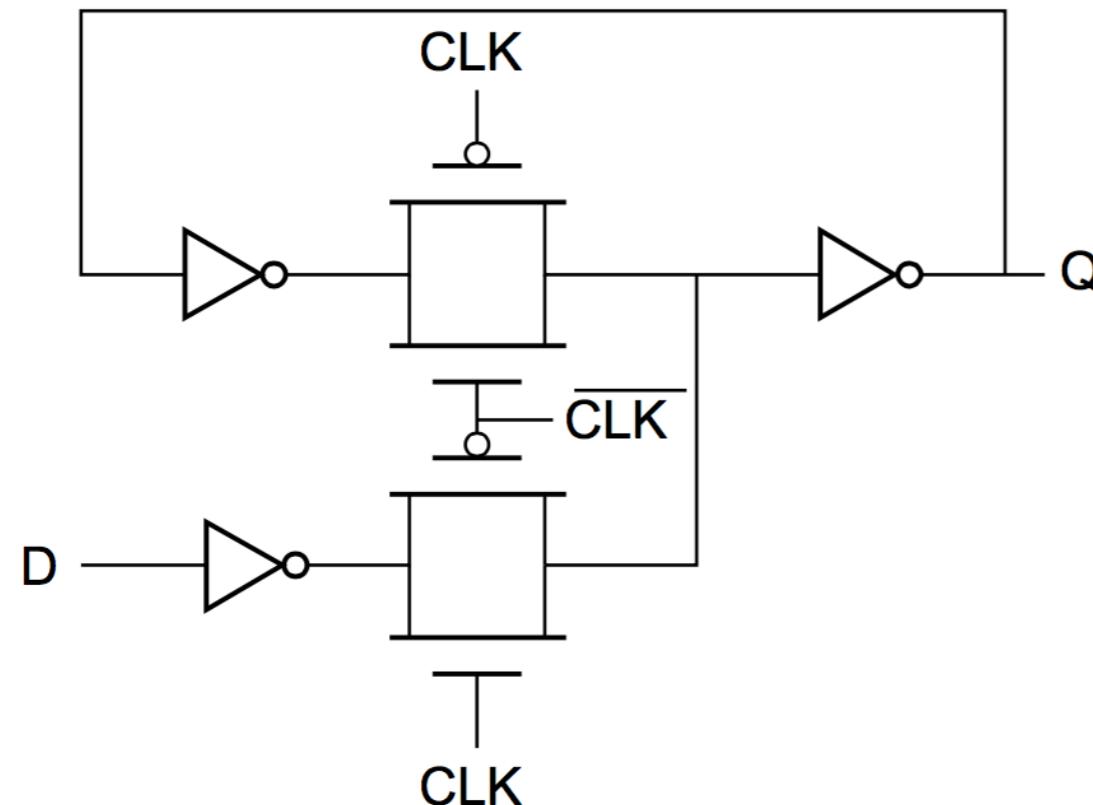


$$Q = D \cdot \overline{CLK} + Q \cdot CLK$$

$$Q = Q \cdot \overline{CLK} + D \cdot CLK$$

Writing into the latch: opening the feedback loop

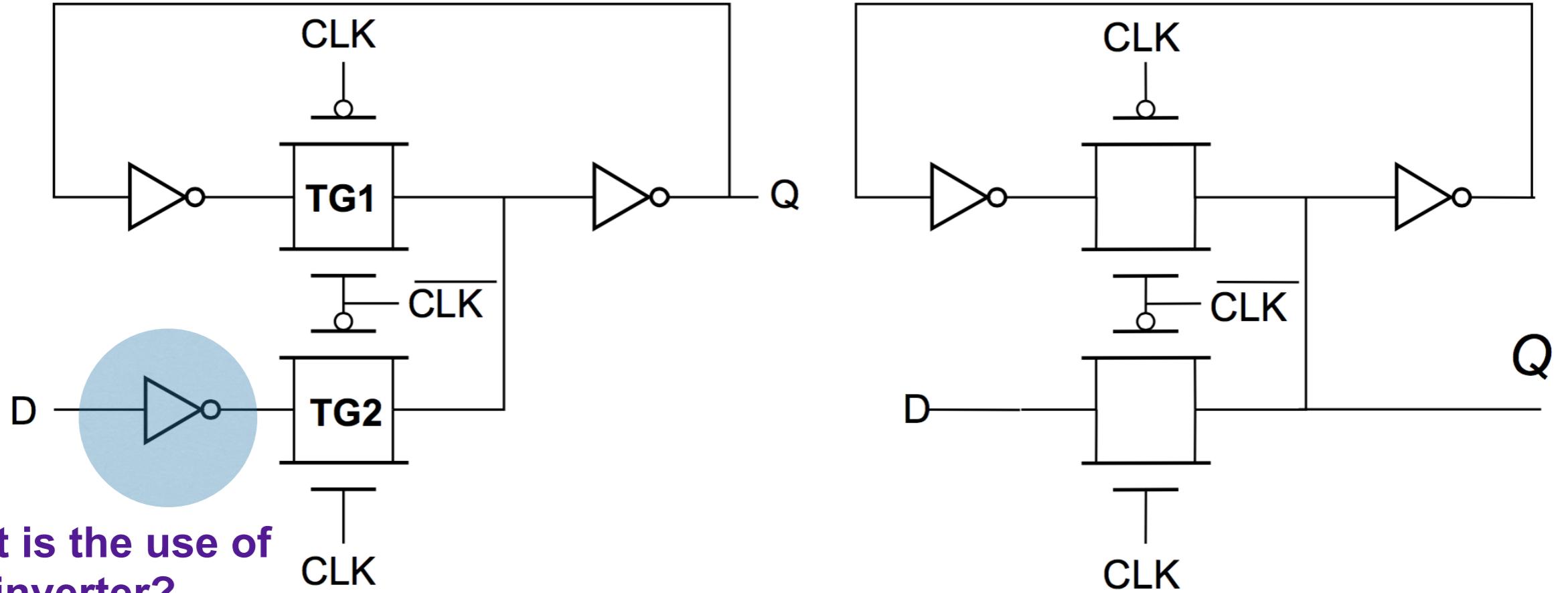
Use the clock as a decoupling signal,
that distinguishes between the transparent and opaque states



$$Q = \overline{CLK} \cdot Q + CLK \cdot D$$

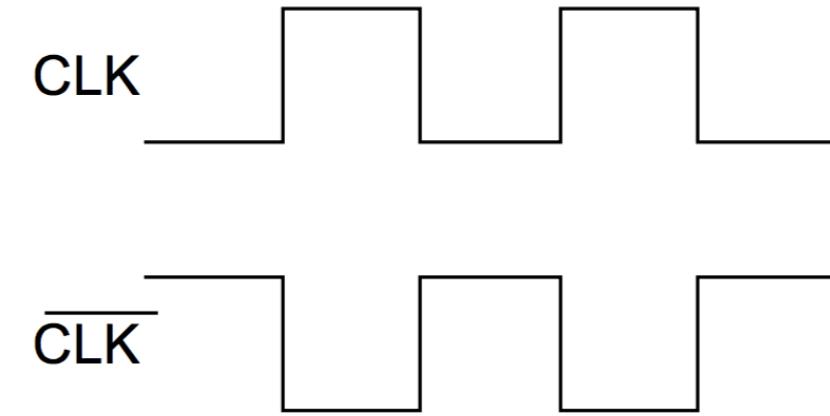
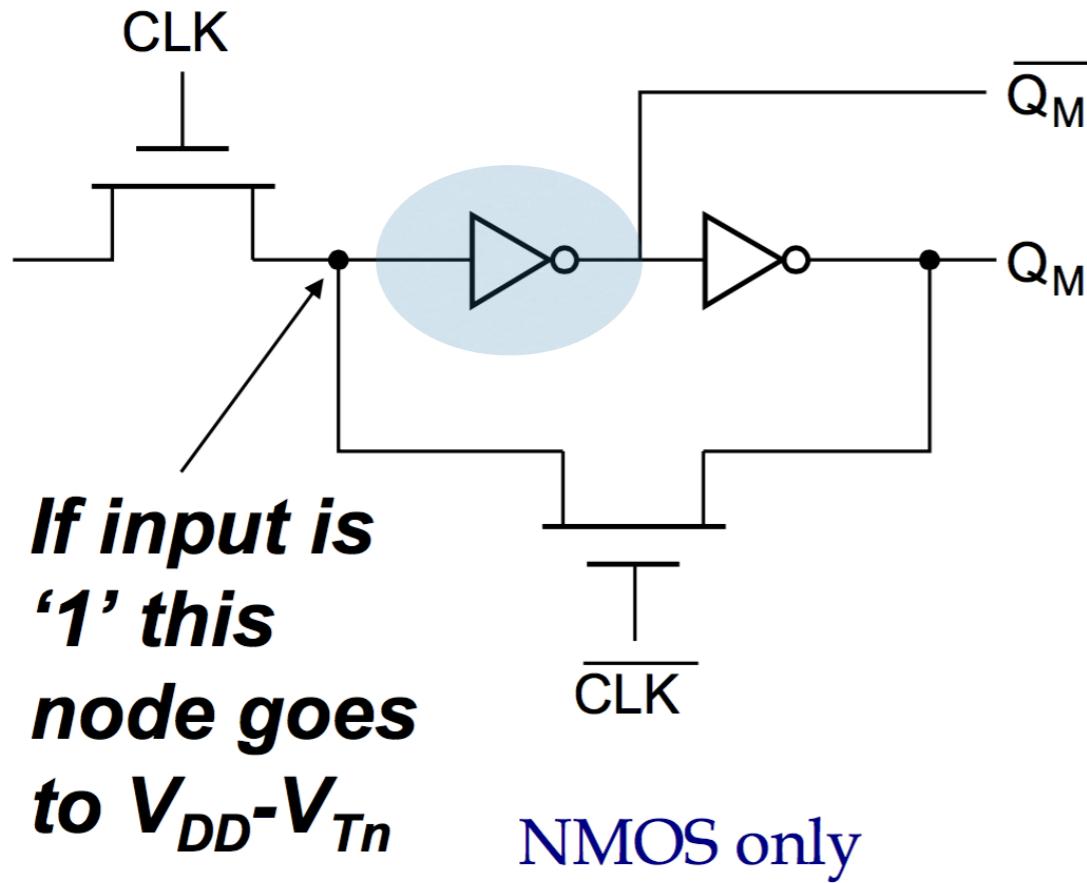
Load on the clock signal is four transistors.

MUX-based latch



- 1) $CLK=0 \Rightarrow TG1$ is ON \Rightarrow Hold state
- 2) $CLK=1 \Rightarrow TG1$ is OFF \Rightarrow Feedback loop open
 \Rightarrow sample the input

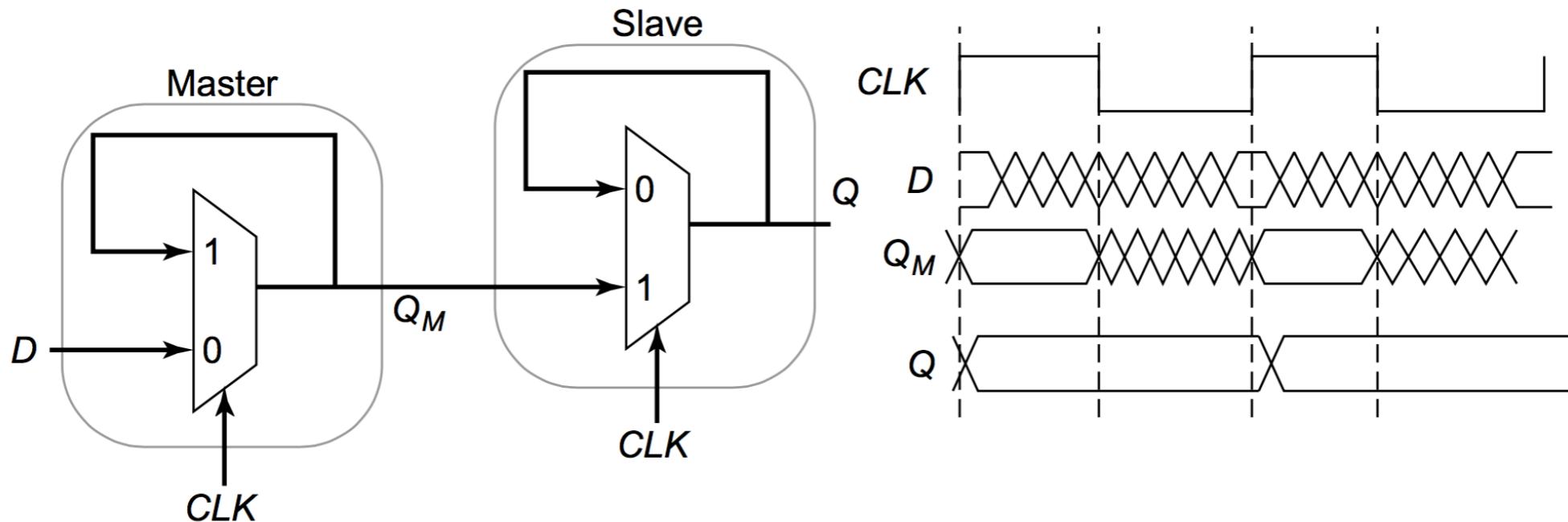
NMOS-only latch



Non-overlapping clocks

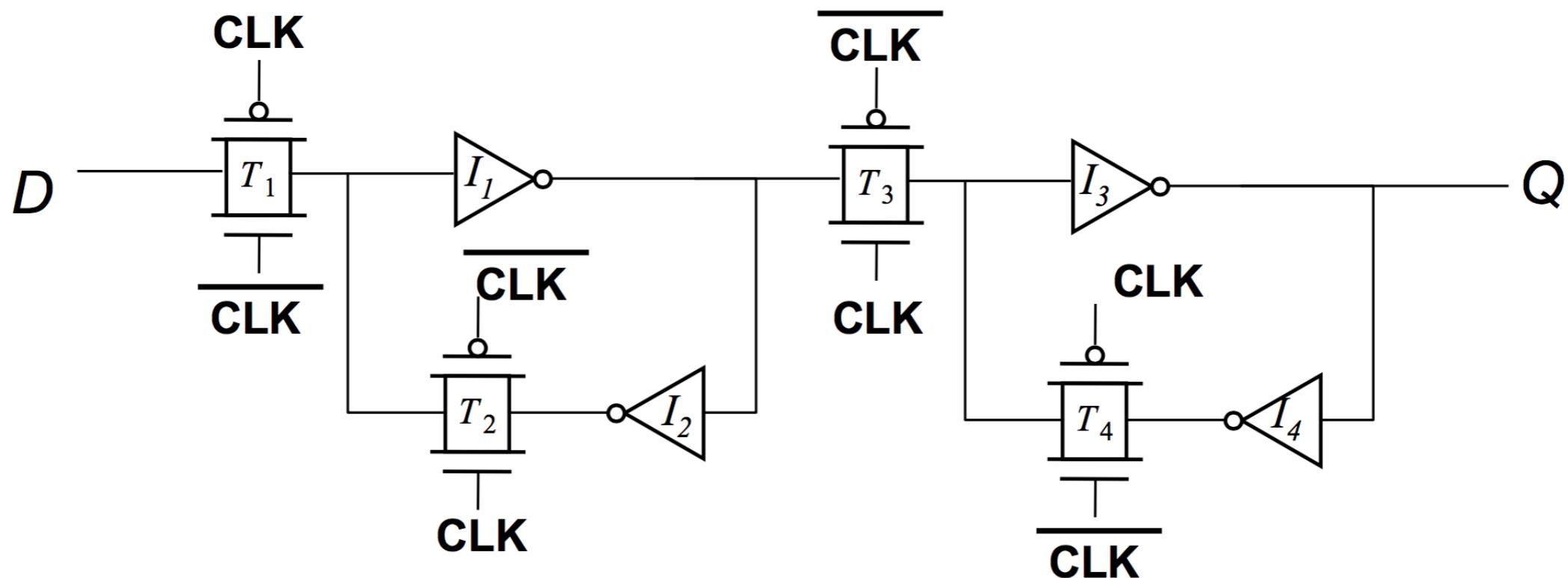
**Reduces the clock load.
Static power in the inverter**

Master-slave register (edge-triggered)



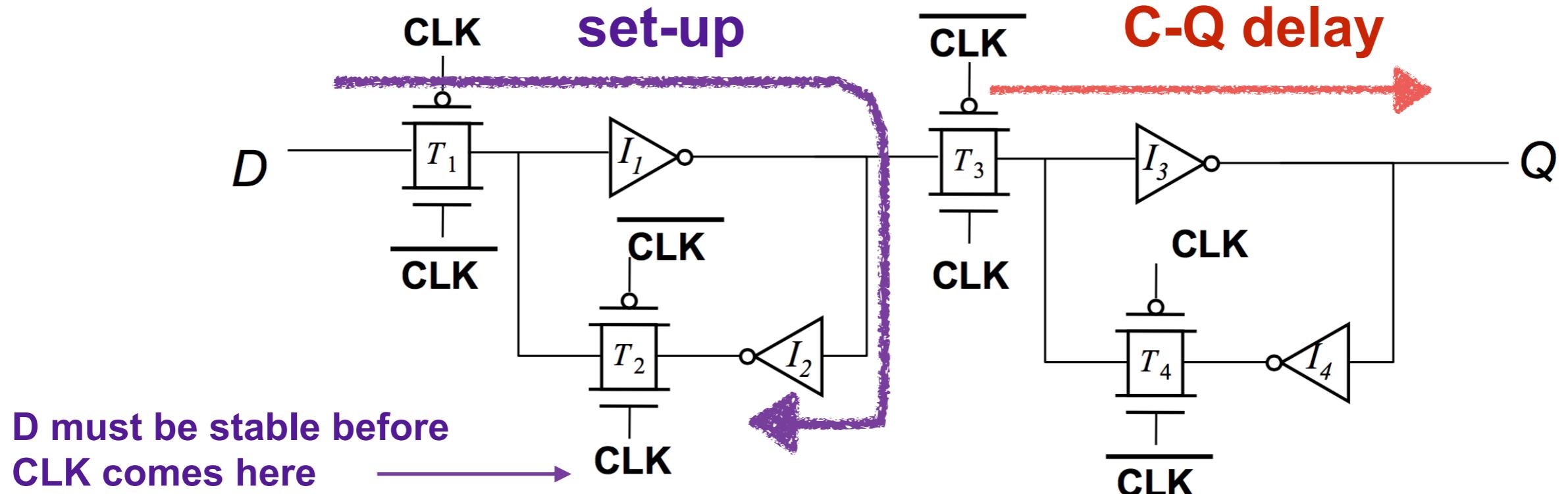
- **Cascading a negative latch and a positive latch**
- **Clock Low => Master is transparent and Slave in Hold**
 - D is passed to Q_M
- **Clock High => Master in hold and Slave is transparent**
 - Q_M is passed to Q (Q_M is constant)
 - Q is same as D just before the rising edge of CLK
 - After CLK is high, changes in D does not modify Q
- **Positive edge-triggered register or flip-flop**

Master-slave flip-flop or register



Find:
Set-up time
C-Q delay
Hold time

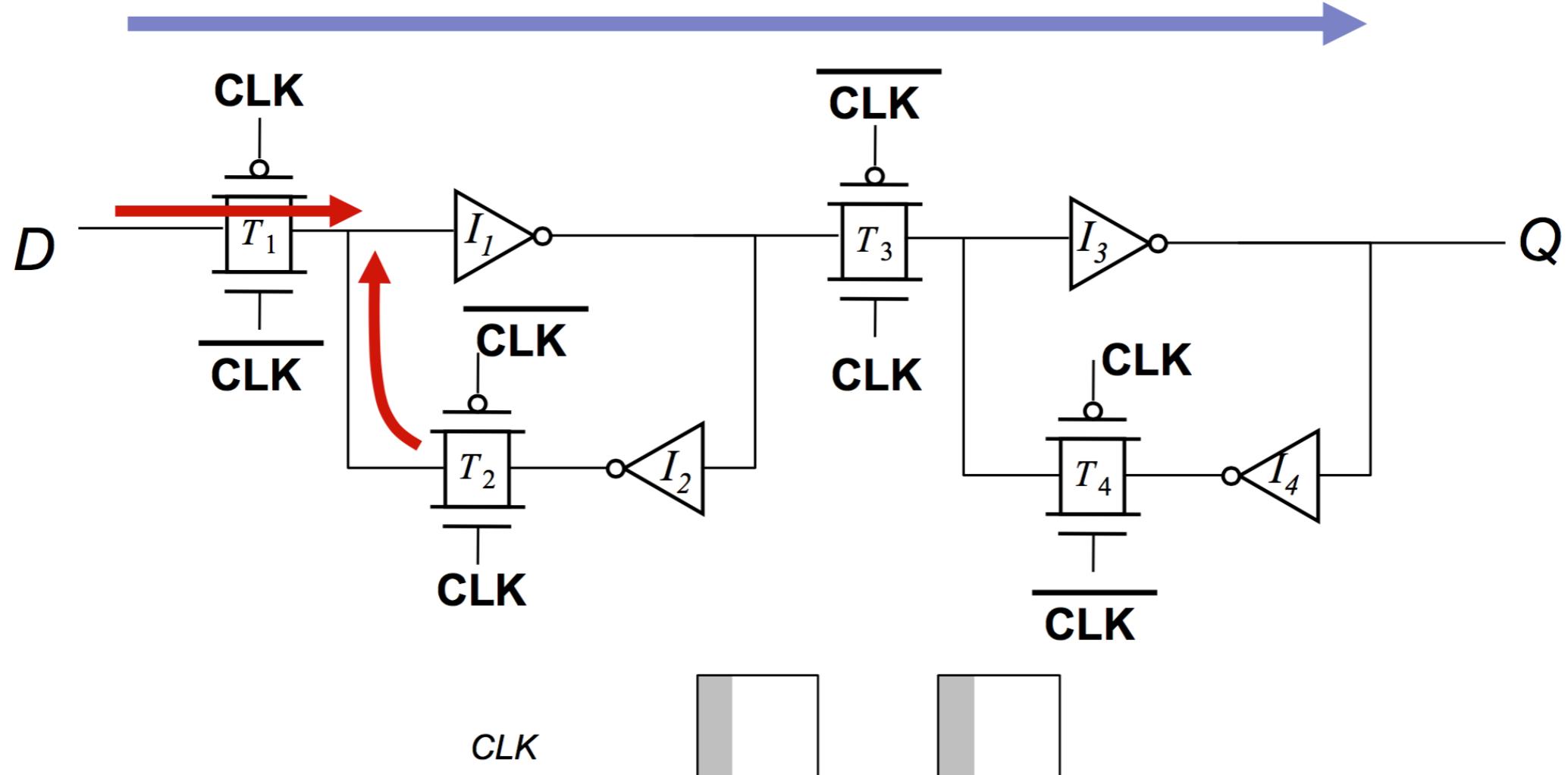
Master-slave flip-flop or register



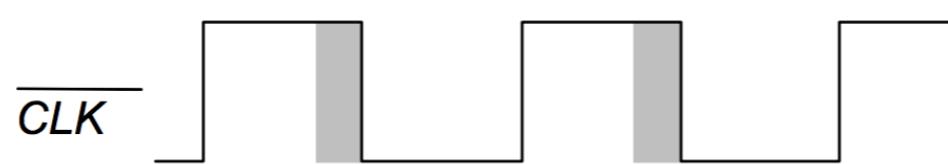
- **Setup time**
 - $T_{\text{setup}} = \text{delay of } T_1 + \text{delay of } I_1 + \text{delay of } I_2$
- **Clock-Q Delay:** $t_{c-q} = \text{delay of } T_3 + \text{delay of } I_3$
- **Hold time:** After clock goes high, T_1 is off $\Rightarrow D$ is isolated from $Q_M \Rightarrow$ hold time is '0'

Clock overlap

Race condition



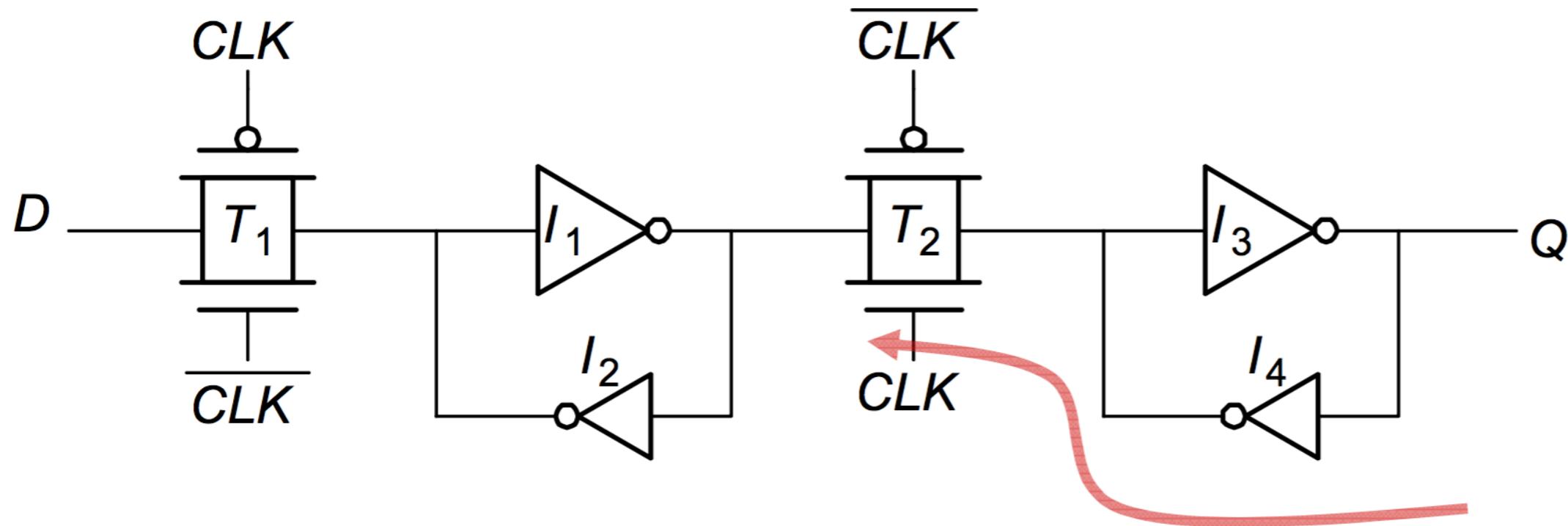
Clock overlap also
called “skew”



clock overlap

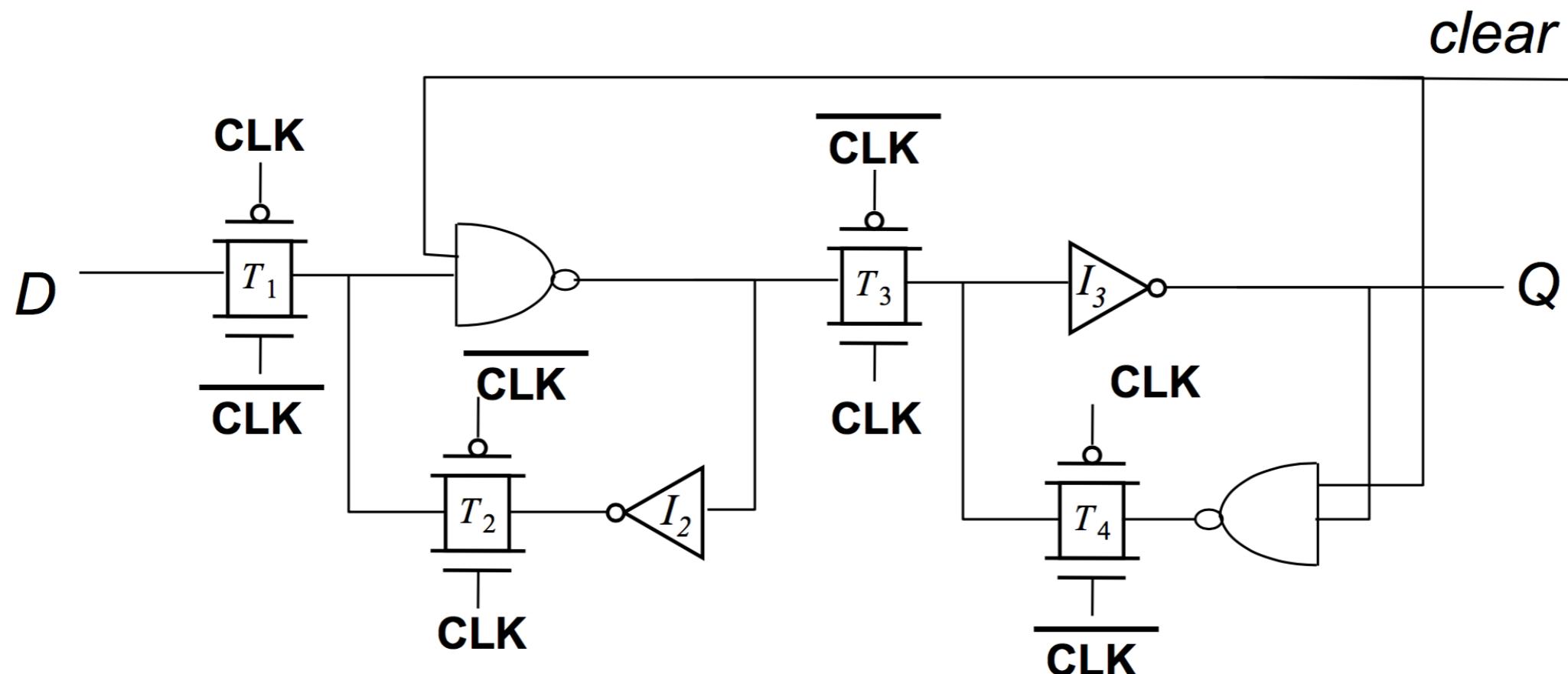
Solution: Use two-
phase clocks

Master-slave flip-flop with reduced clock load



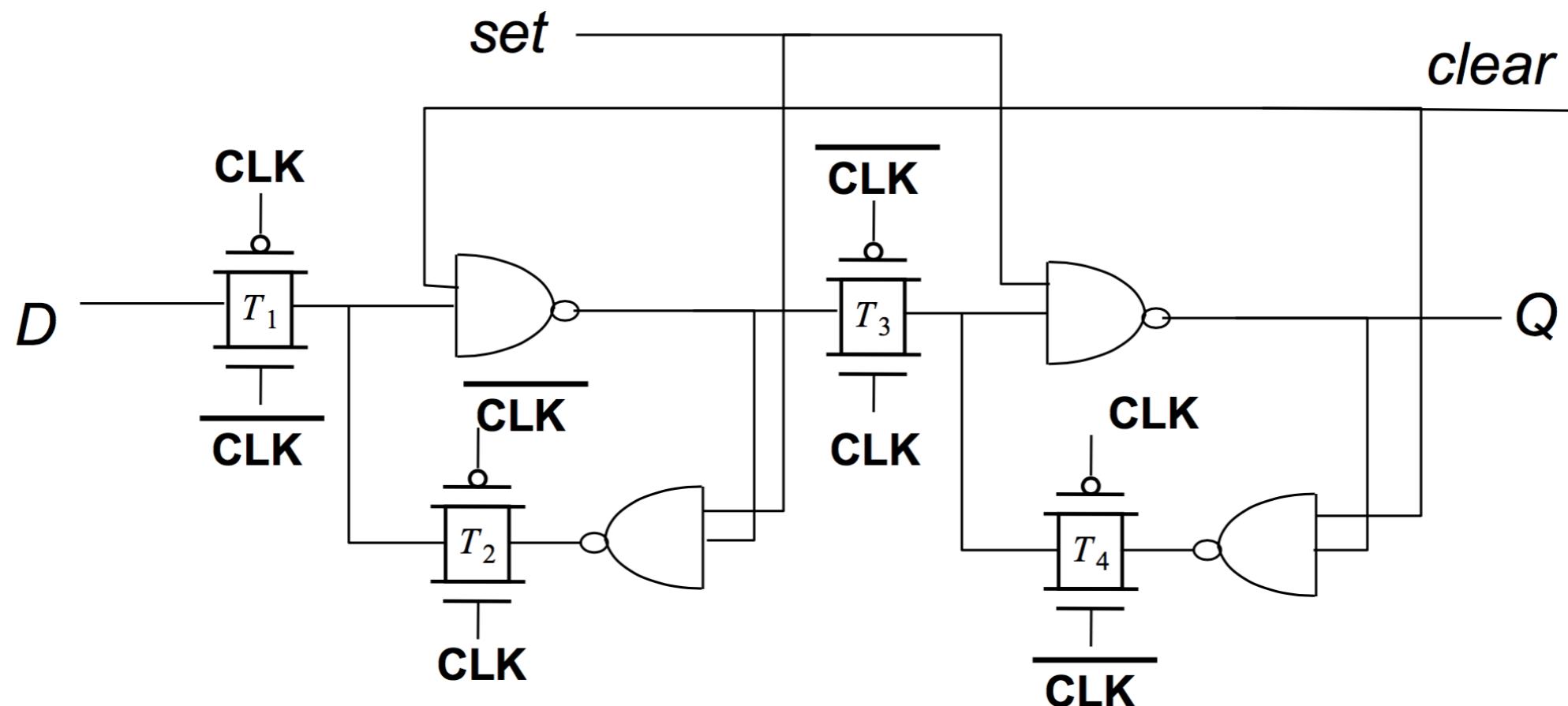
- **Lower clock load**
- **Complex design:**
 - T_1 and driver of D need to overpower I_2
 - I_2 needs to be designed weak
 - Reverse conduction is possible when slave is ON

Master-slave flip-flop with clear



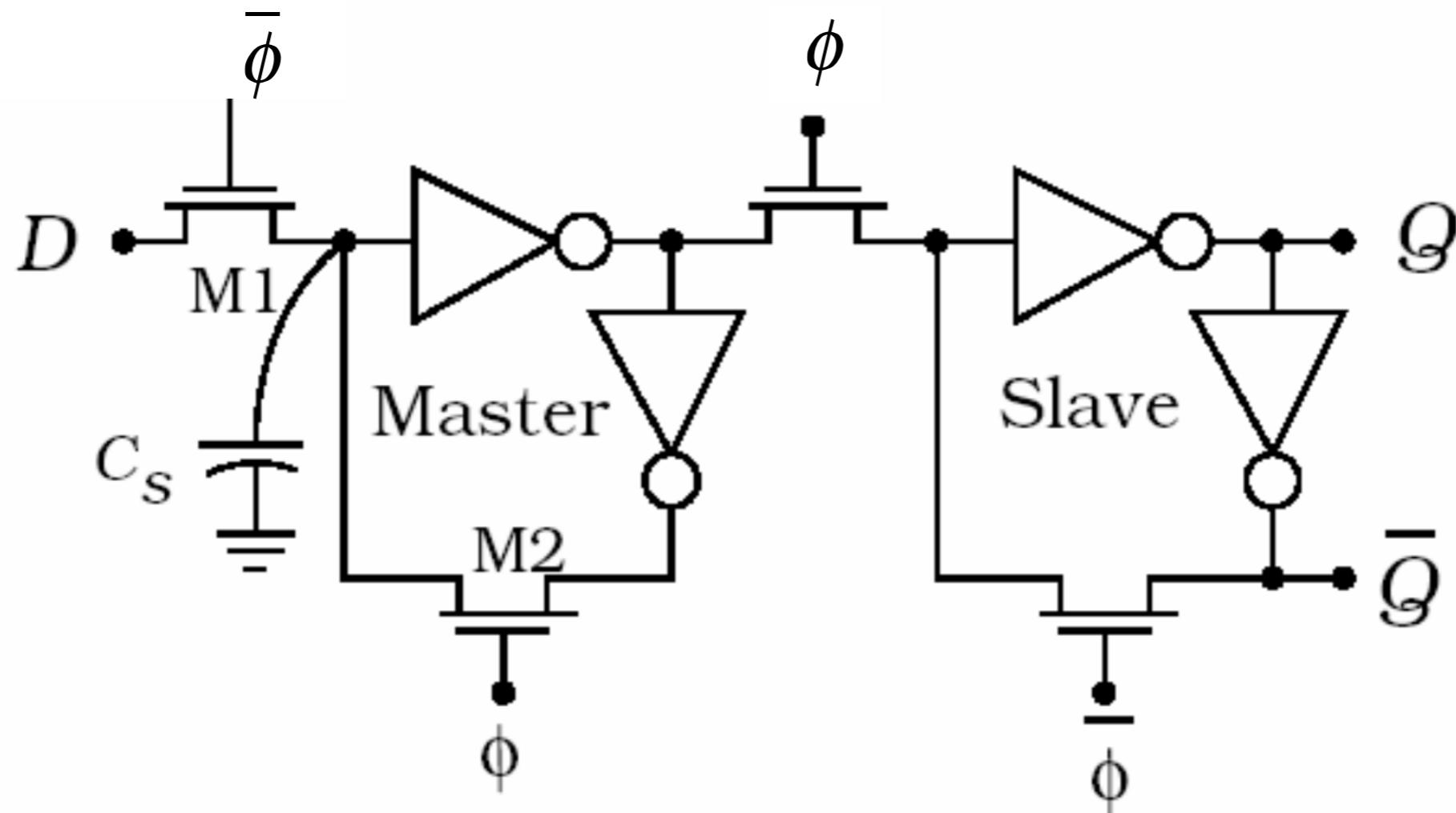
$clear = 1 \Rightarrow$ NAND gates act as inverter \Rightarrow regular DFF
 $clear = 0 \Rightarrow Q$ is forced to '0'

Master-slave flip-flop with clear and set



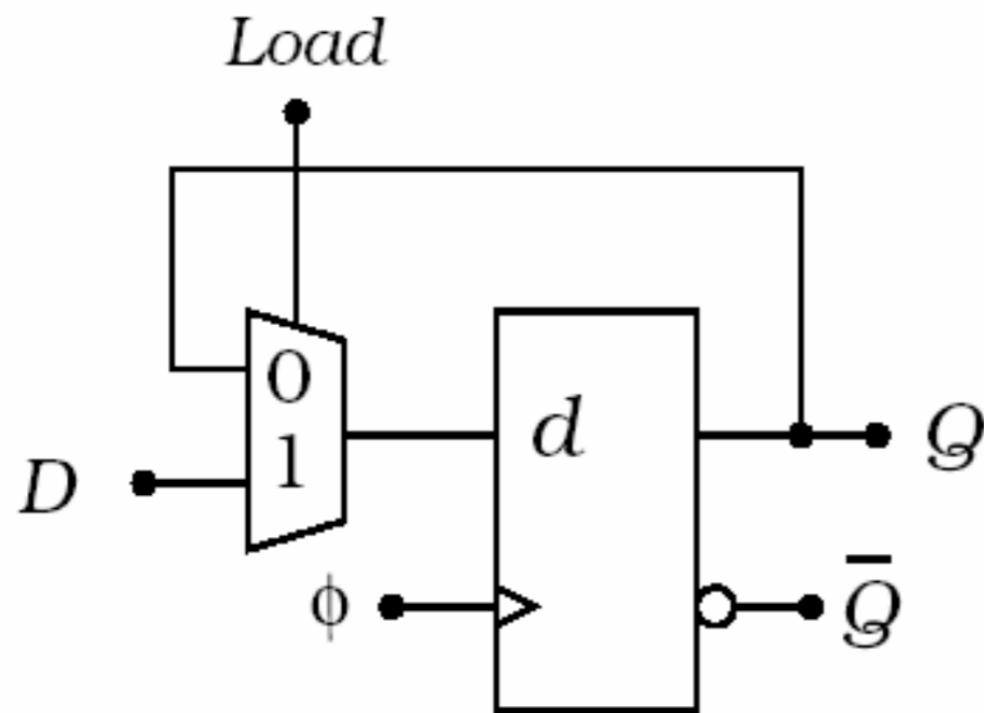
Set	Clear	Operation
1	1	Regular DFF
0	1	$Q = 1$
1	0	$Q = 0$
0	0	Forbidden

Master-slave flip-flop with pass transistors

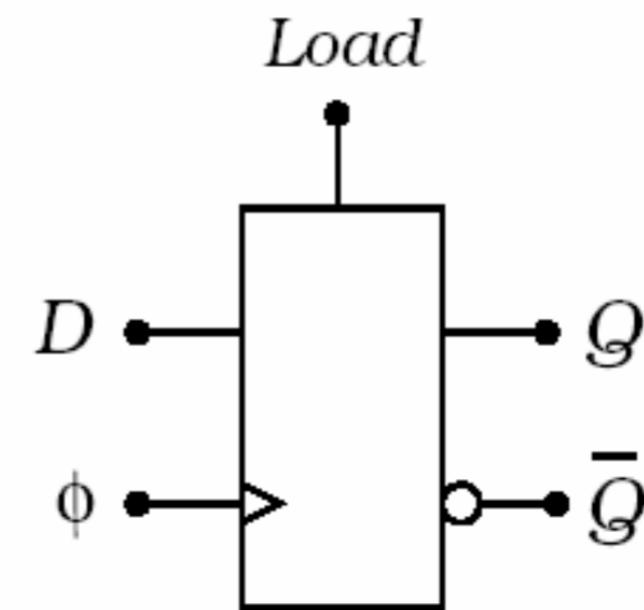


- Reduces clock load to FOUR transistors
- Static power dissipation in inverter
- Also has the same clock overlap (race condition) issues that TG-based D flip flop has

D flip-flop with load-control



(a) Wiring diagram

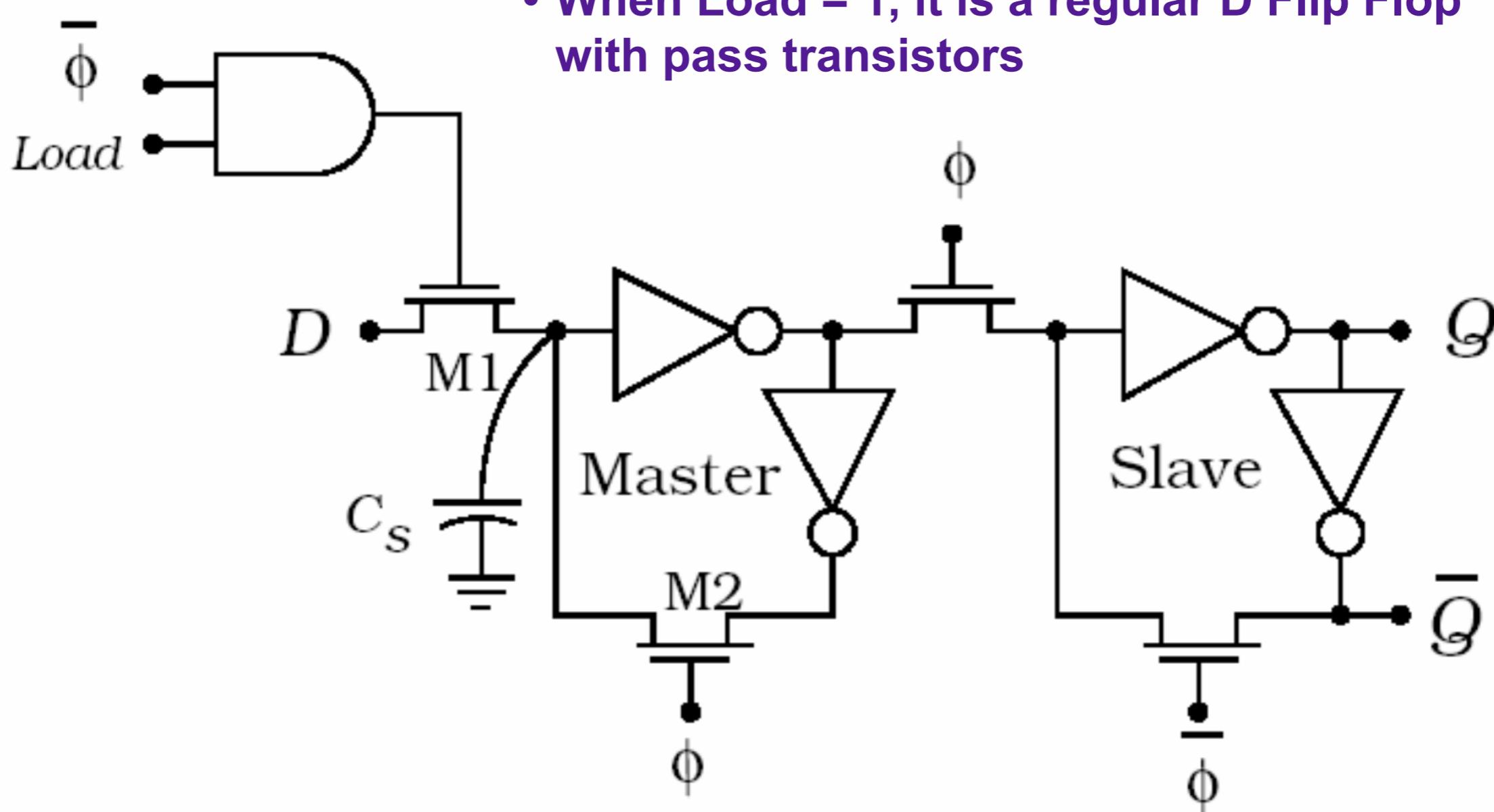


(b) Symbol

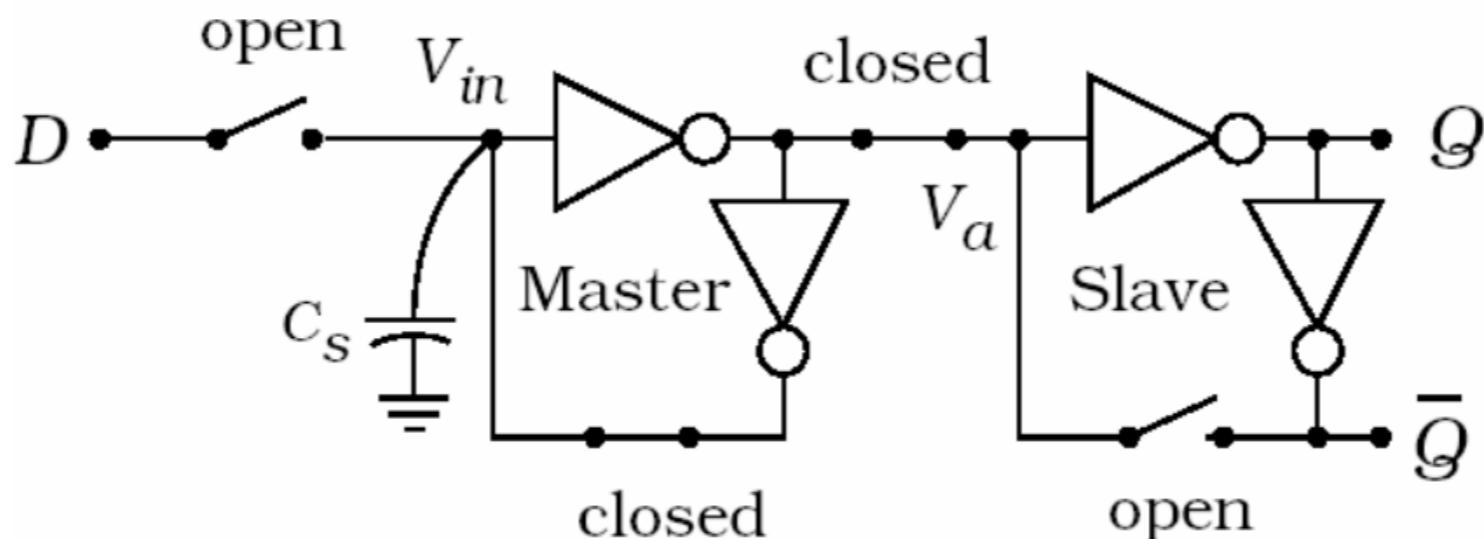
Load the input when Load = 1

Store the data when Load = 0

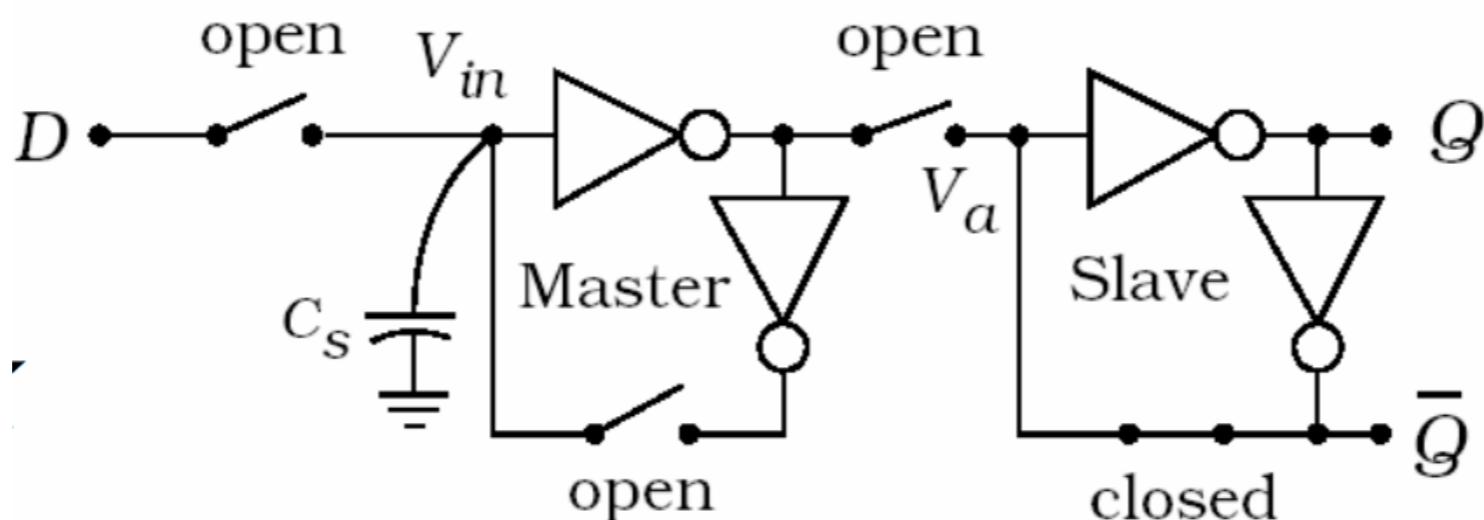
Master-slave flip-flop with load-control



Operation for Load = 0

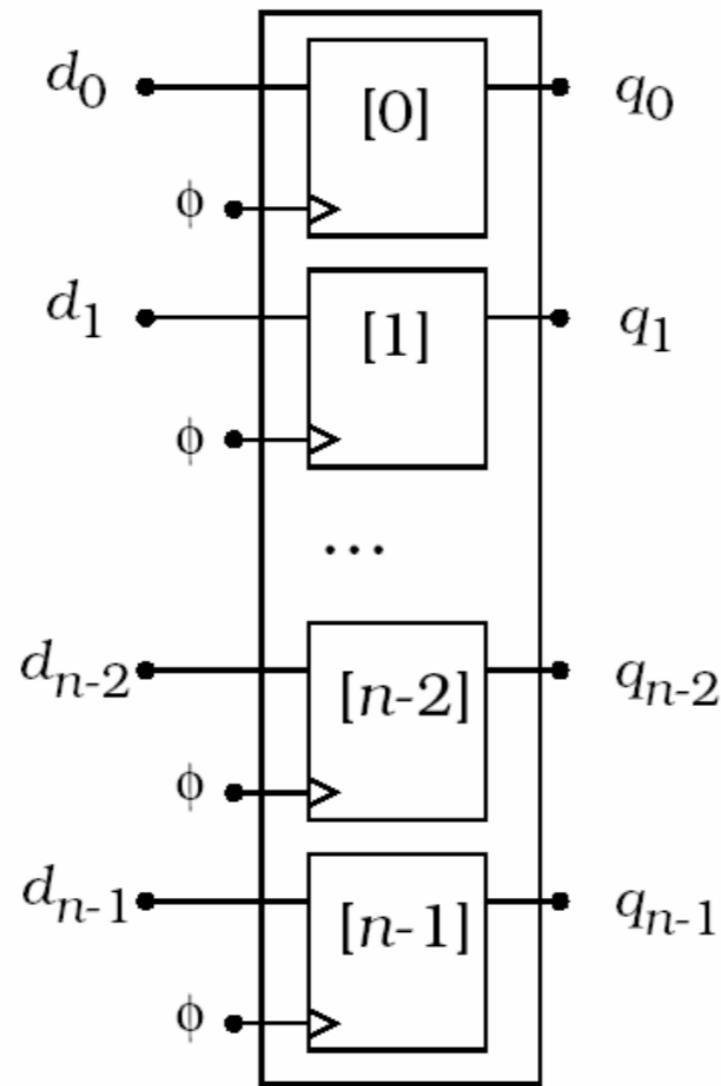


(b) Hold with $\phi = 1$, $\text{Load} = 0$

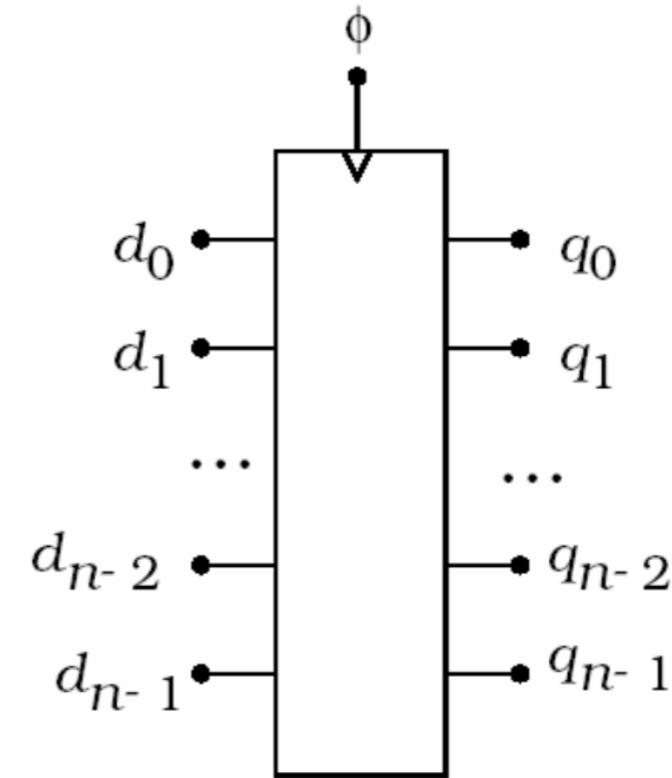


(c) Hold with $\phi = 0$, $\text{Load} = 0$

Construction of n-bit register



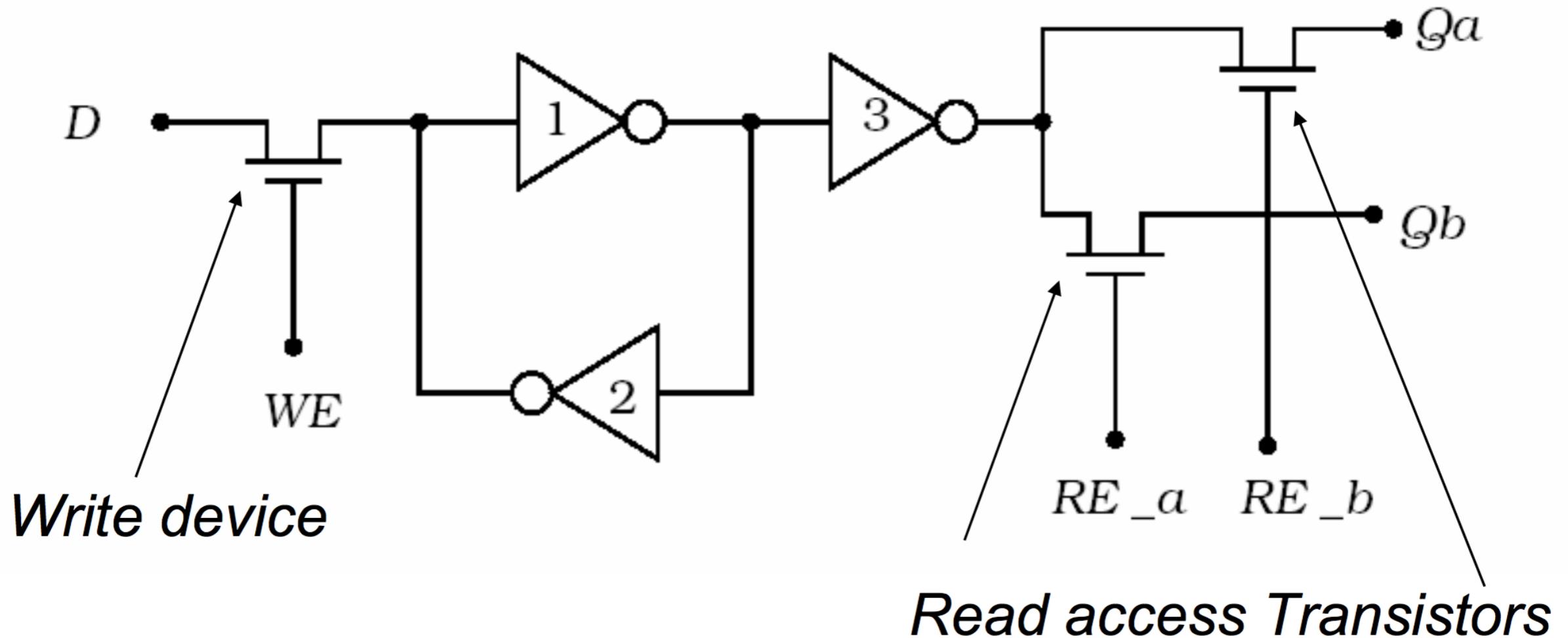
(a) Internal construction



(b) Basic symbol

*Store state values of a logic circuit
Large clock power dissipation*

One-bit static multi-port register



For clocked loading we can use ($WE \cdot CLK$)

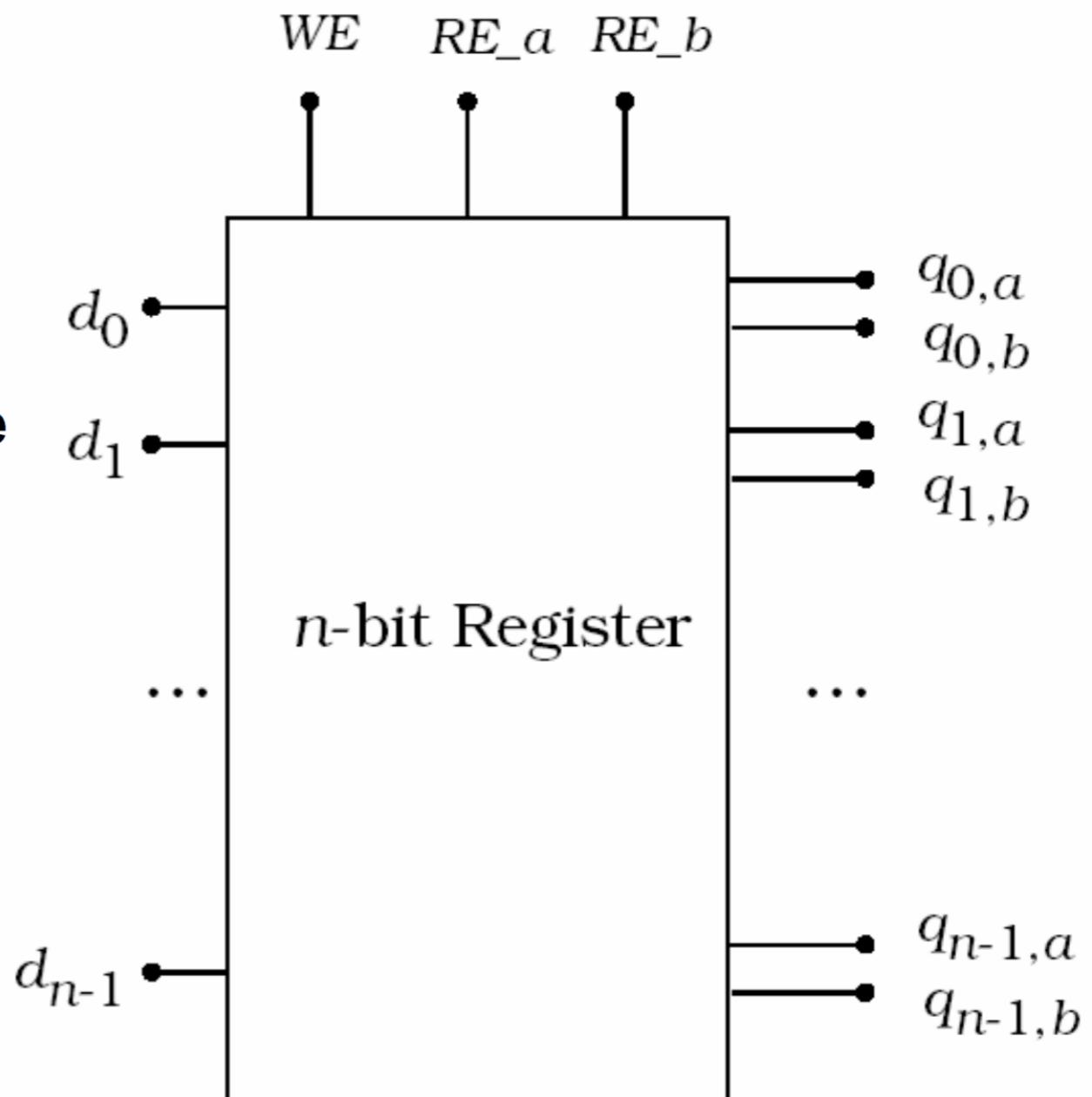
N-bit static multi-port register

Multi-port register file

Clock power is low

Useful when a state needs to be held for a long time.

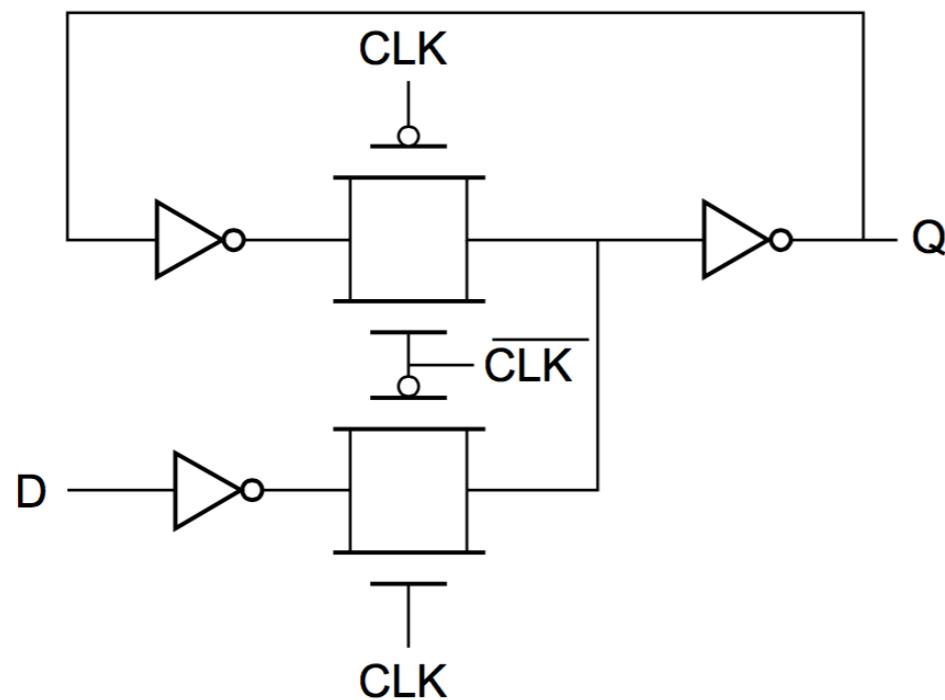
A regular DFF will incur a large clock power dissipation although no useful operations is being performed



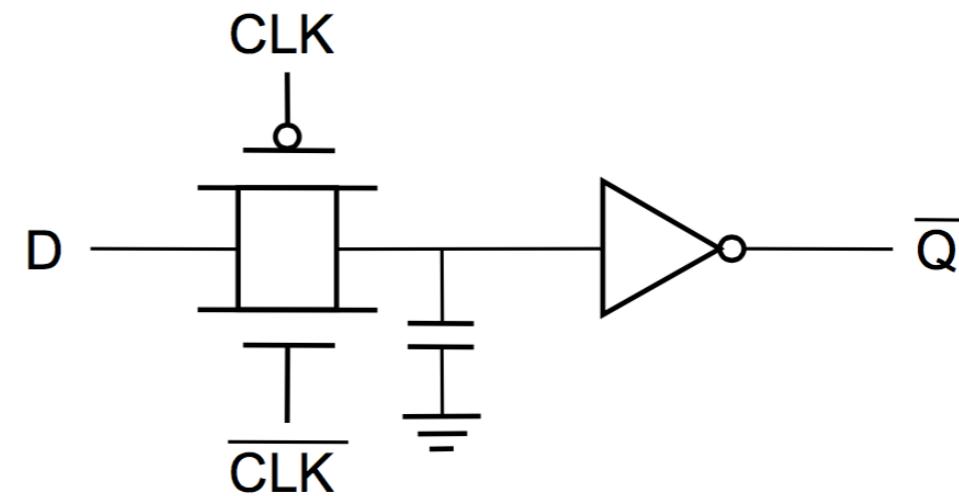
Dynamic latches

Storage mechanisms

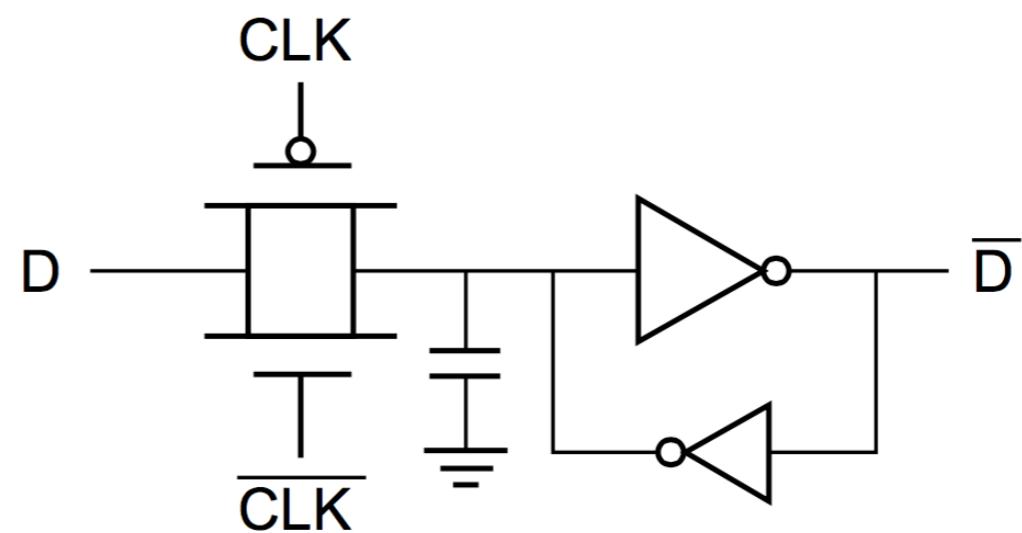
Static



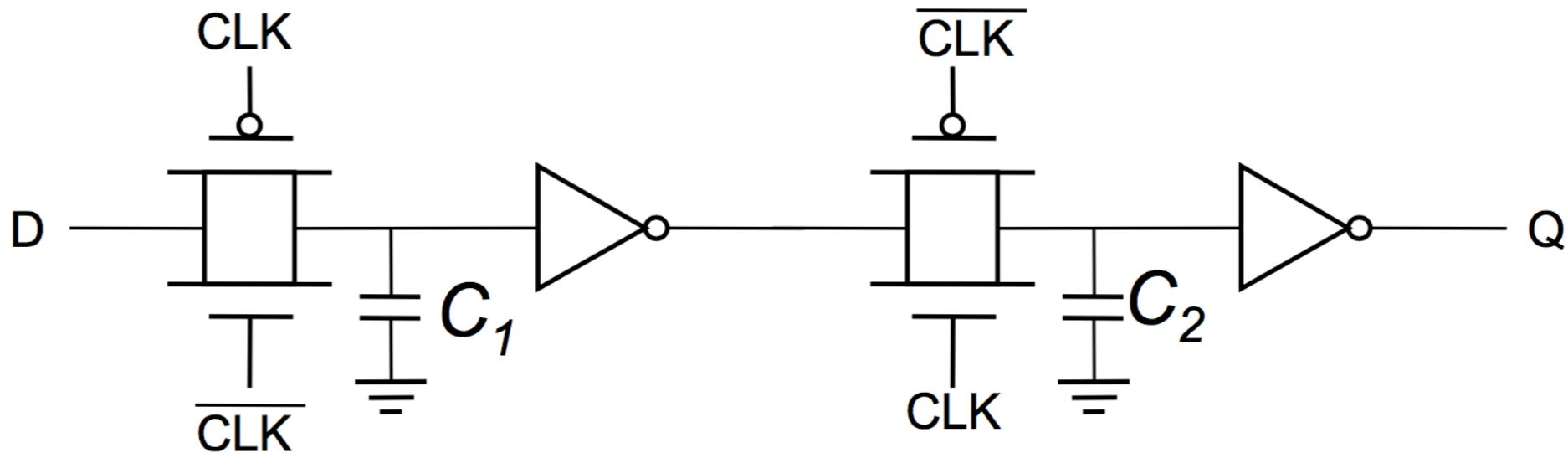
Dynamic (charge-based)



Pseudo-Static



Dynamic master-slave flip flop



$\text{CLK} = 0 \Rightarrow$ sample the input in C_1 ; slave in hold mode (data stored in C_2)

$\text{CLK} = 1 \Rightarrow$ master in hold mode (data stored in C_1)
slave samples the master data and drives Q

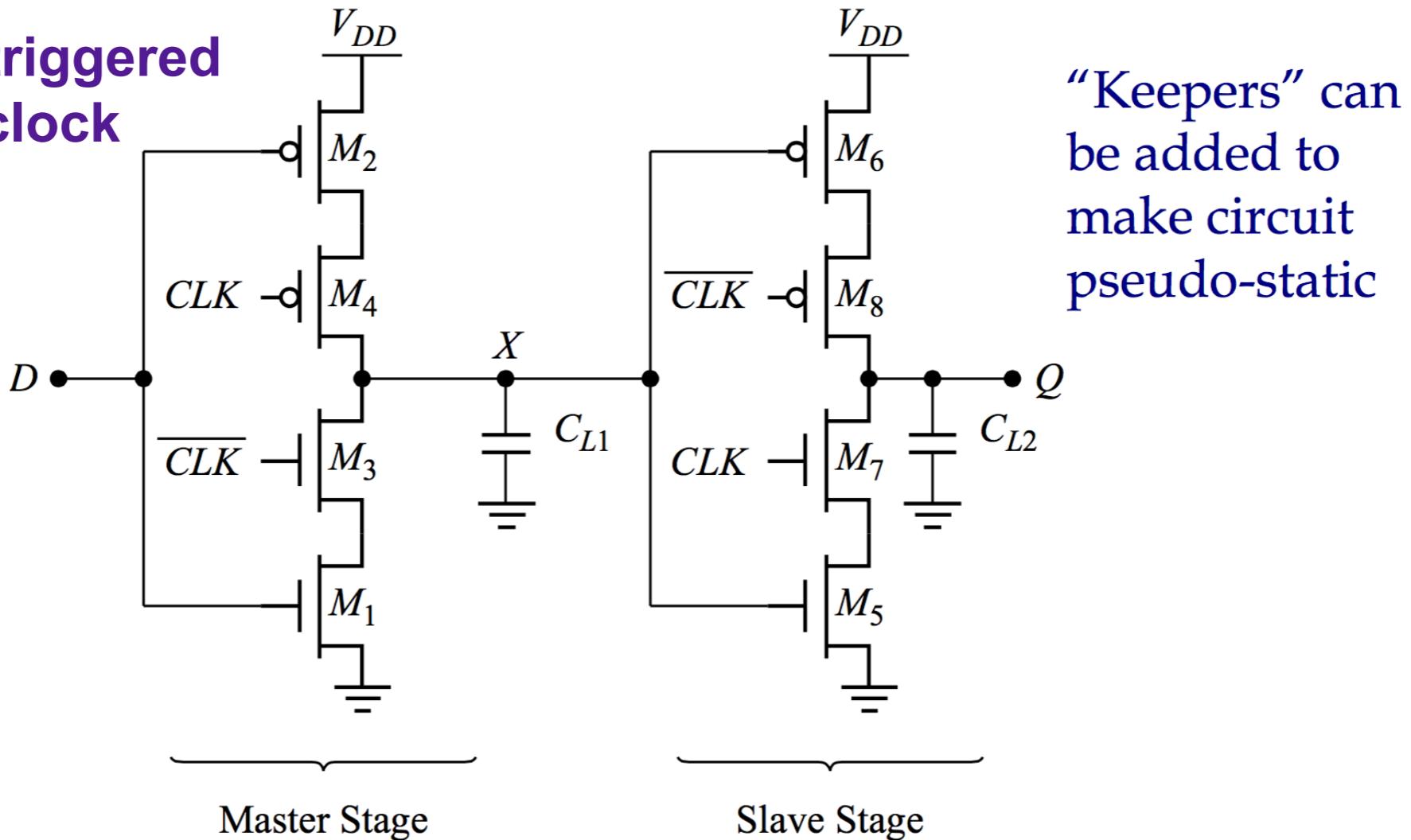
Setup time: delay of first Transmission gate

CLK-Q : delay of second TG + delay of slave inverter

CLK - $\overline{\text{CLK}}$ overlap causes race condition

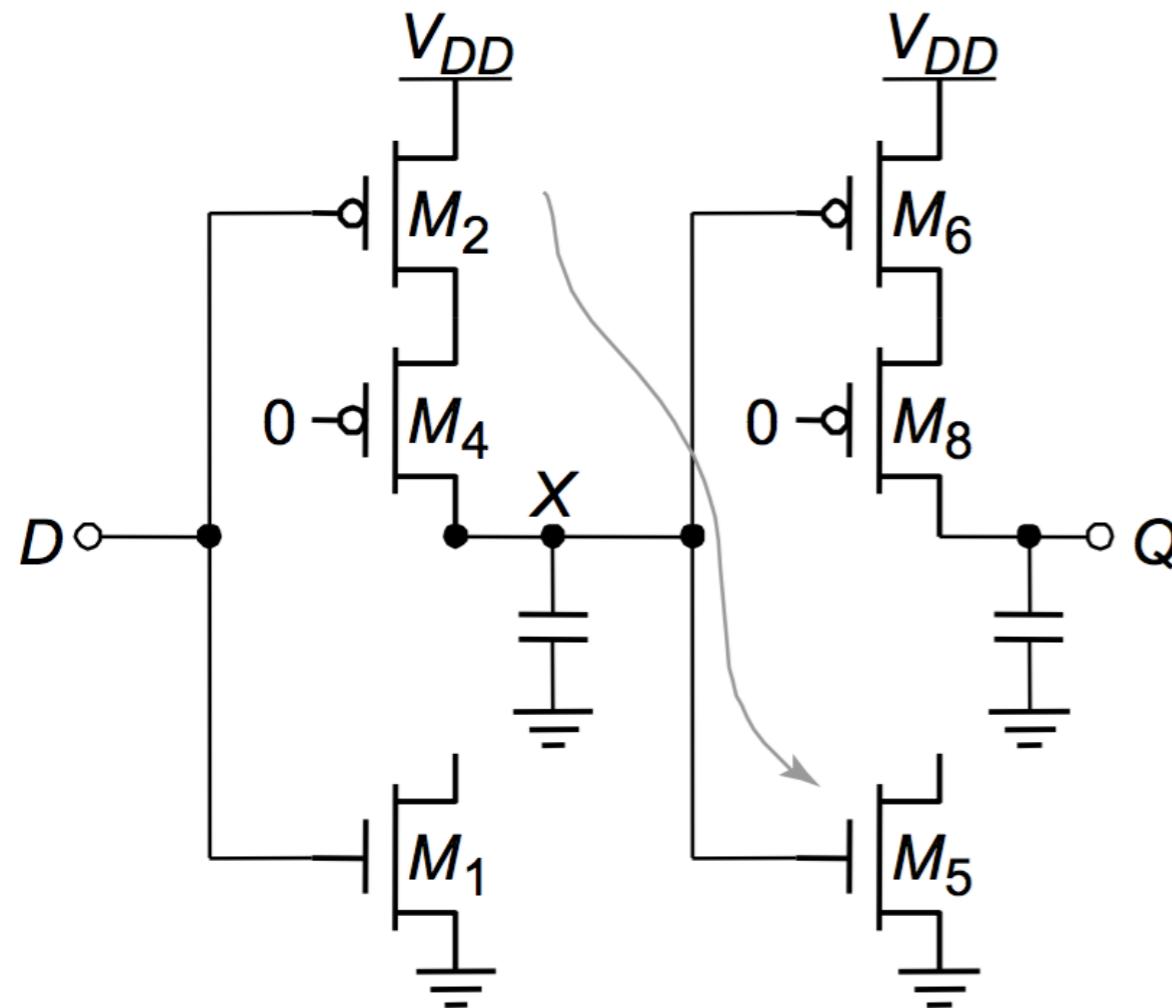
C^2MOS master-slave flip flop

- Positive edge triggered
- Insensitive to clock overlap



$CLK = 0 \Rightarrow D$ is sampled (inverted) to X ; slave in hold
 $CLK = 1 \Rightarrow$ master in hold; X is sampled to Q (inverted)

In insensitive to clock skew



(a) (0-0) overlap: CLK and !CLK are '0'

Q is disconnected from D

CLK and !CLK = 0 for a small time-interval

We need to ensure in this time interval the change in D does not propagate to 'Q'

If D makes a $0 \rightarrow 1$ transition X is disconnected from D

If D makes a $1 \rightarrow 0$ transition, X makes a $0 \rightarrow 1$ transition, but Q cannot be discharged as M_7 is OFF

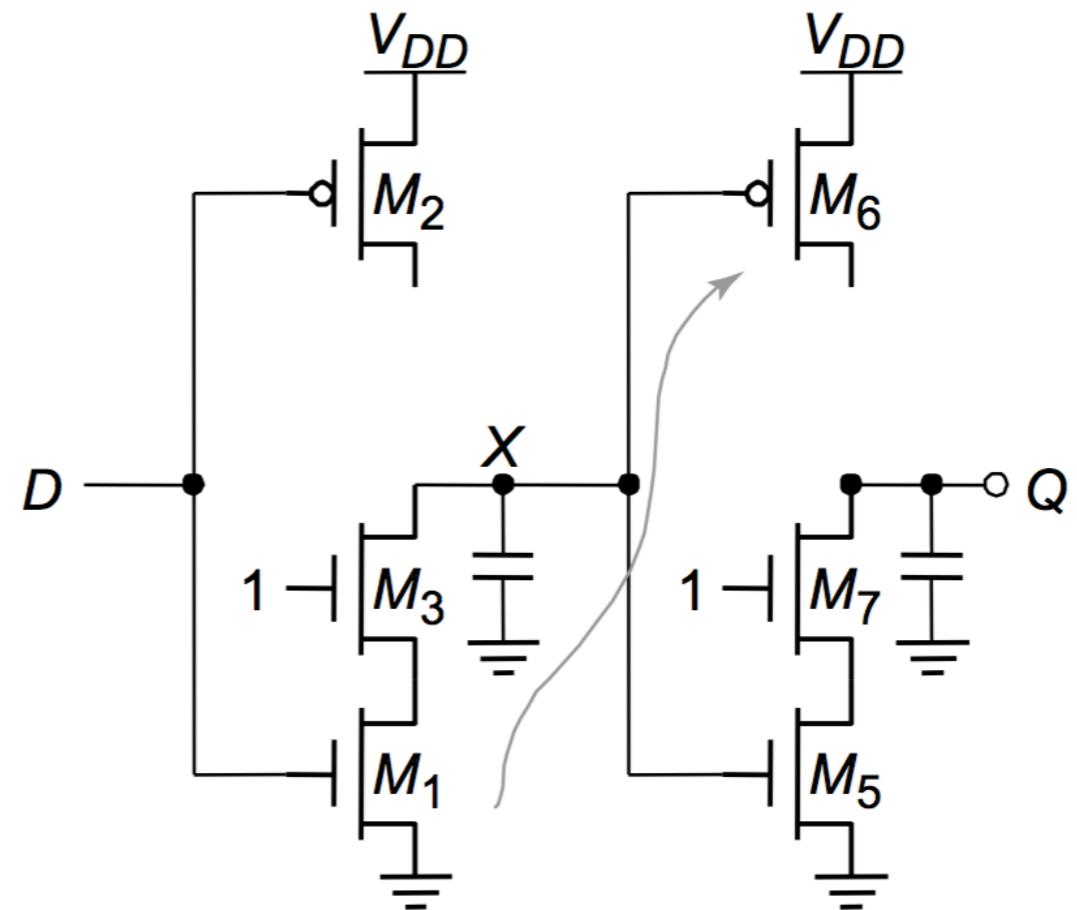
Insensitive to clock skew

CLK and !CLK = 1 for a small time-interval

We need to ensure in this time interval the change in D does not propagate to 'Q'

If D makes a $0 \rightarrow 1$ transition, X makes a $1 \rightarrow 0$ transition, but Q cannot be charged as M_8 is OFF

If D makes a $1 \rightarrow 0$ transition, X is disconnected from D as M_4 is OFF.



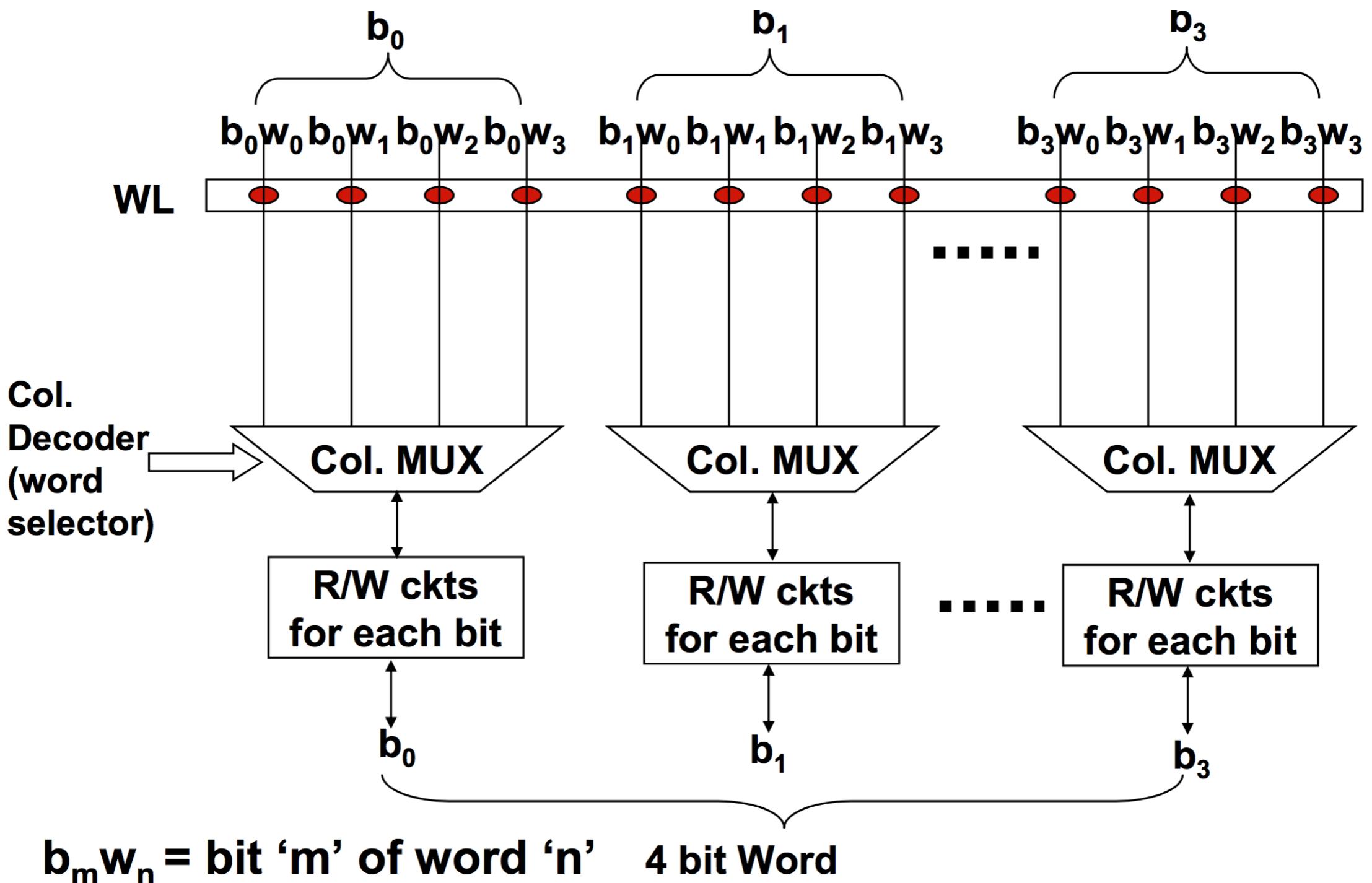
(b) (1-1) overlap: CLK and !CLK are '0

Q is disconnected from D

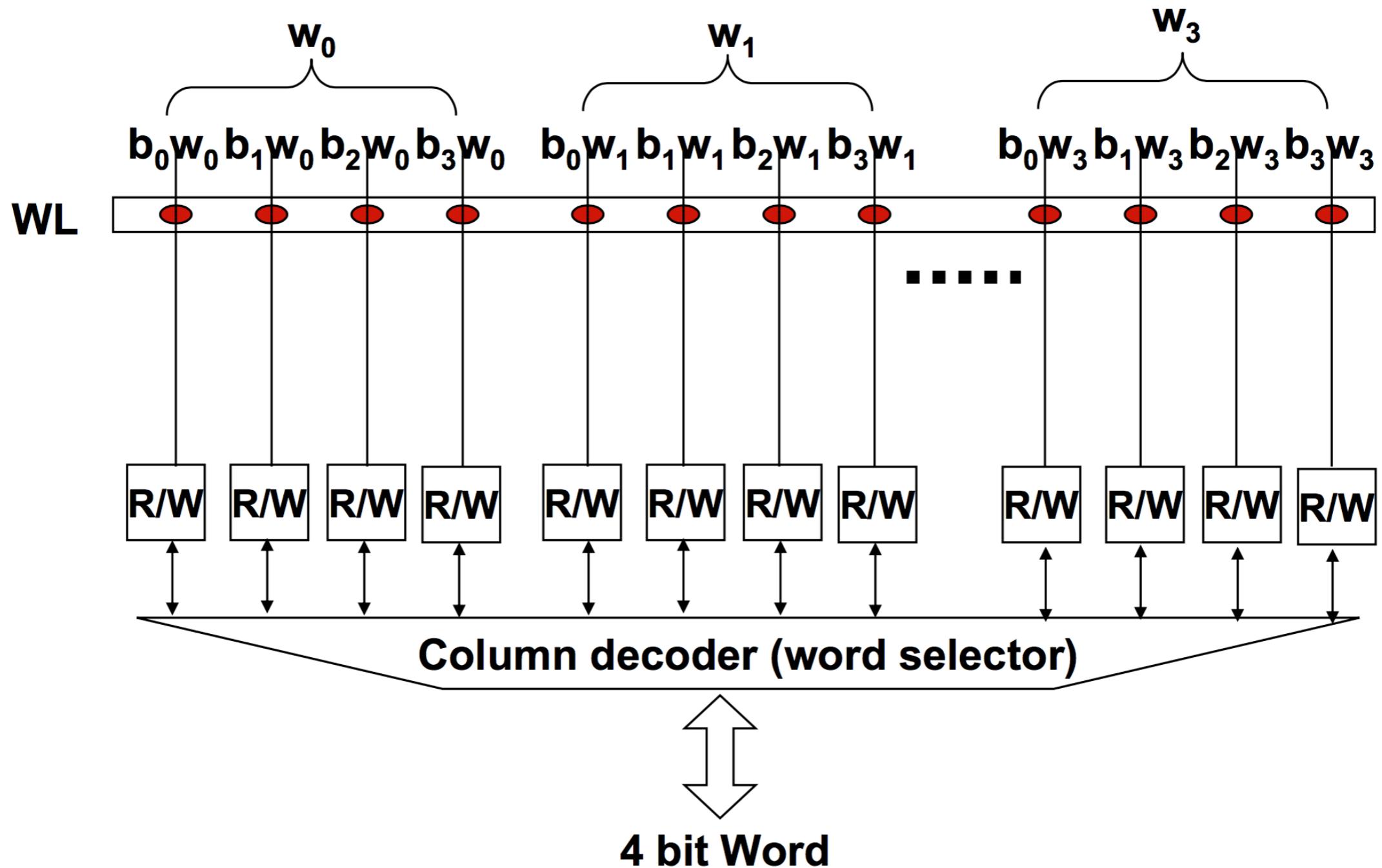
Memories

(not on the mid-term)

Array-structured memory arch.

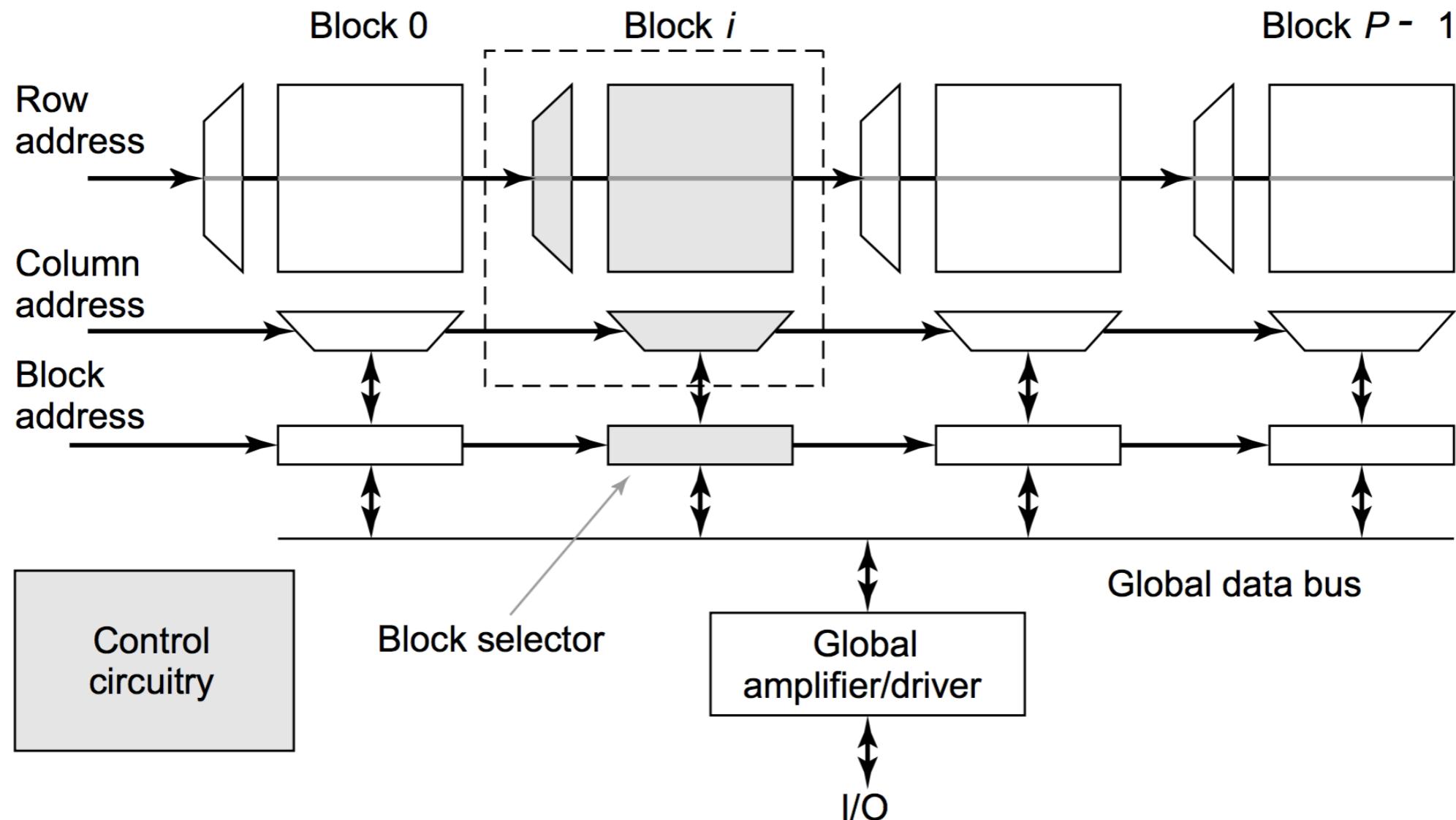


Array-structured memory arch.



$b_m w_n = \text{bit } 'm' \text{ of word } 'n'$

Array-structured memory arch.



Advantages:

1. Shorter wires within blocks
2. Block address activates only 1 block => power savings

Memory types

□ STATIC (SRAM)

Data stored as long as supply is applied

Large (6 transistors/cell)

Fast

Differential

□ DYNAMIC (DRAM)

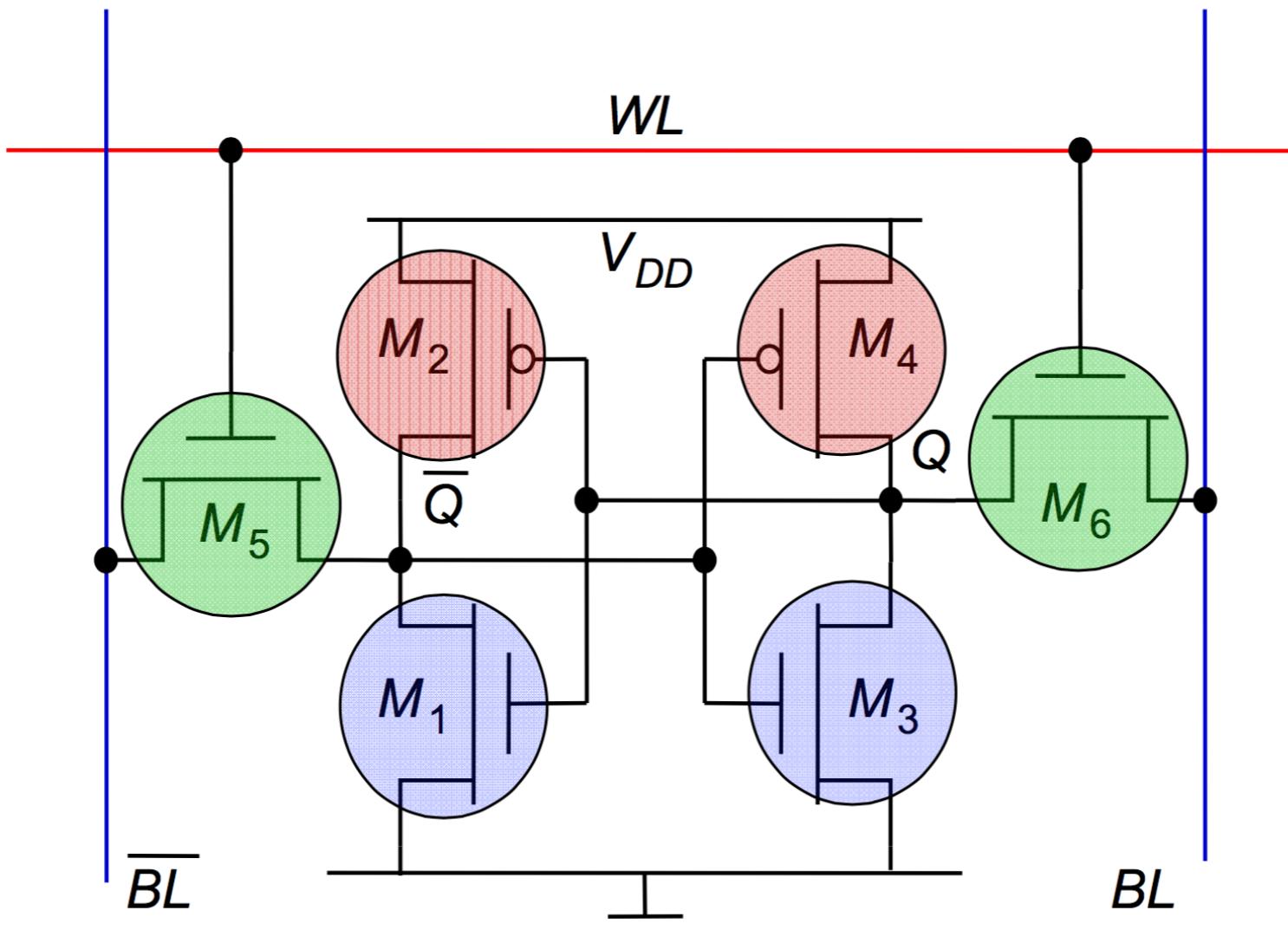
Periodic refresh required

Small (1-3 transistors/cell)

Slower

Single Ended

6T CMOS SRAM cell



● pull-up or load devices (M₂, M₄)

● pull-down devices (M₁, M₃)

● access or pass-gate devices (M₅, M₆)

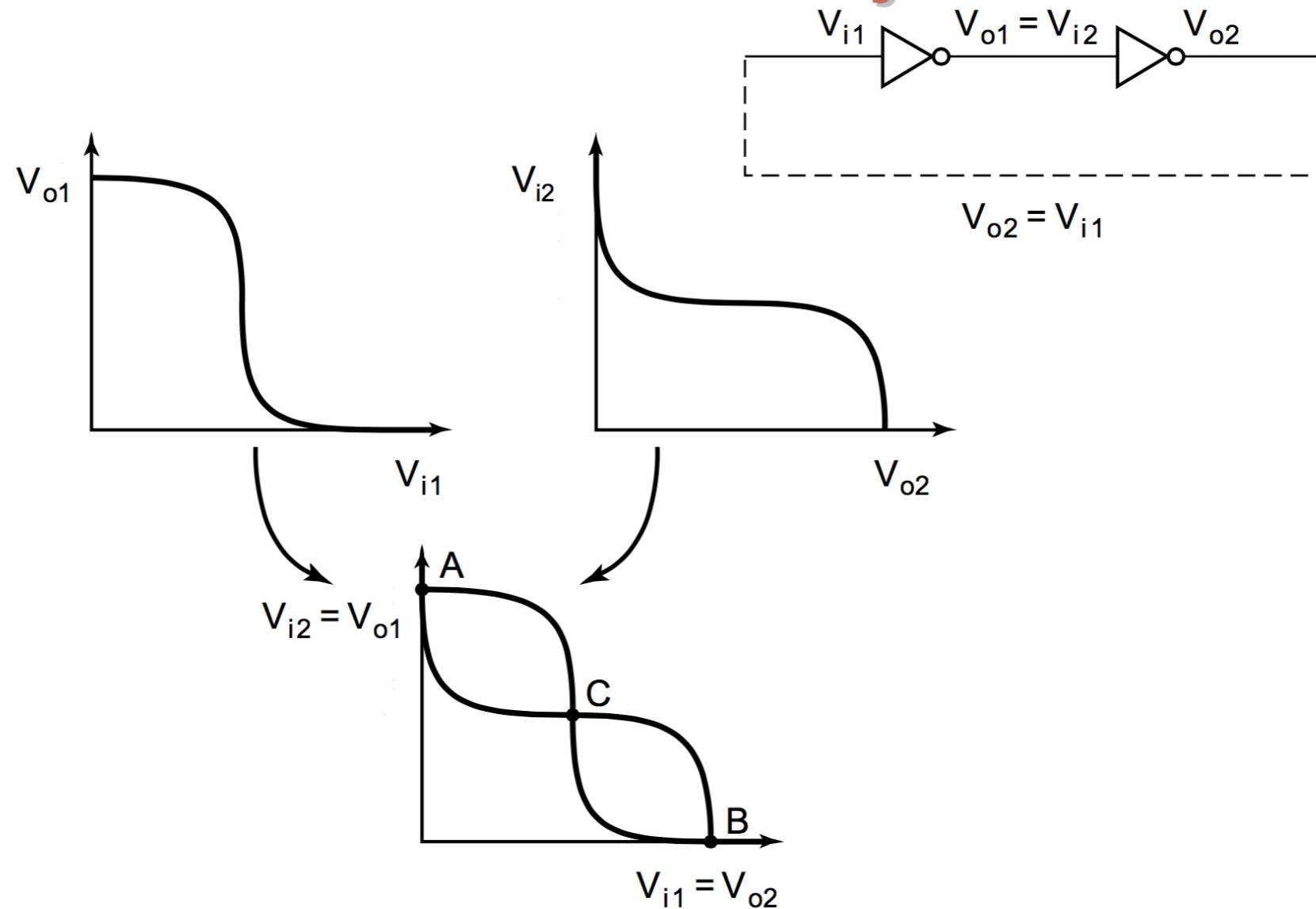
□ Data Holding

□ Read

□ Write

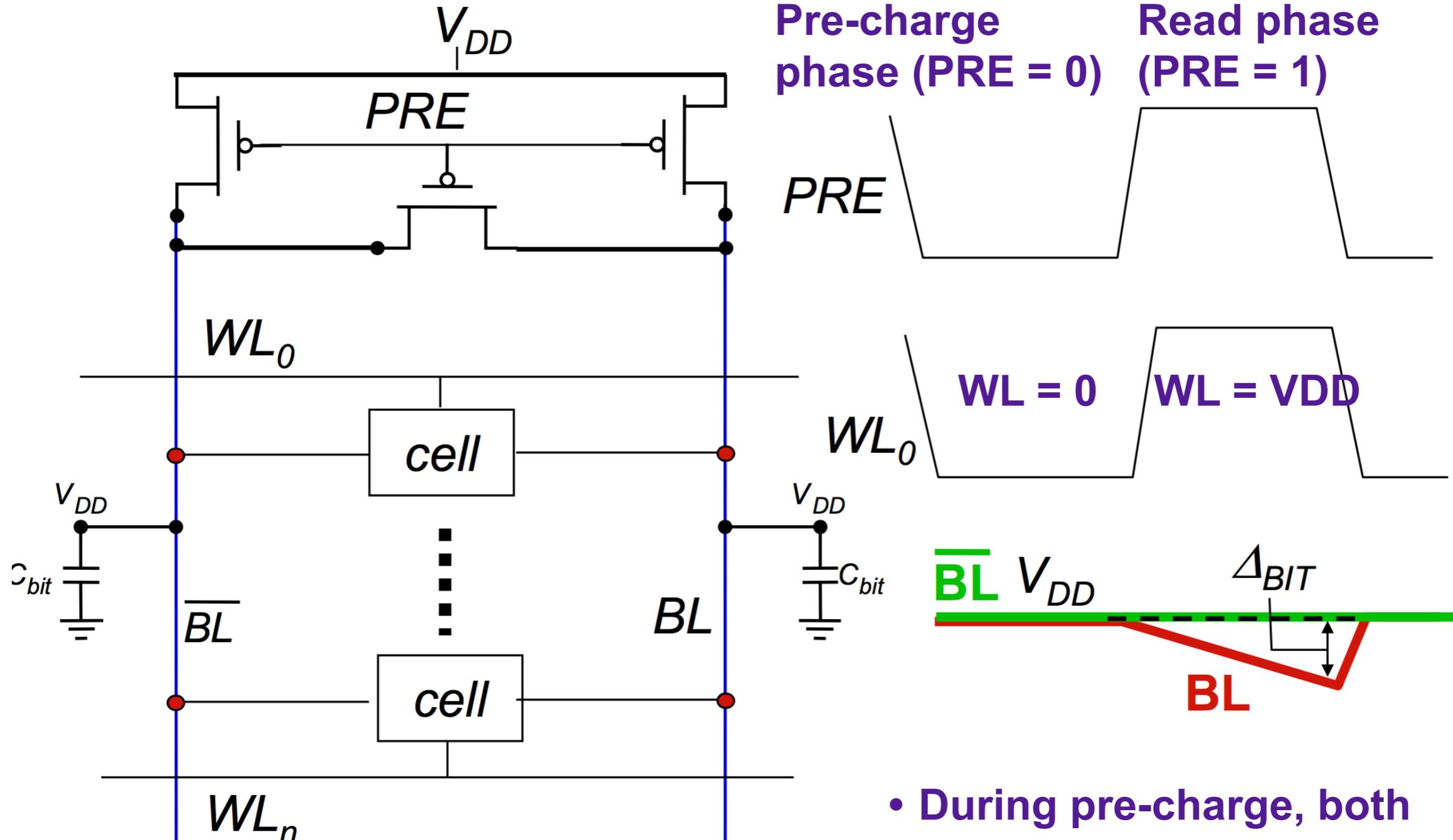
Basic principle

Positive Feedback: Bi-Stability



A cross-coupled inverter provides a basic storage element

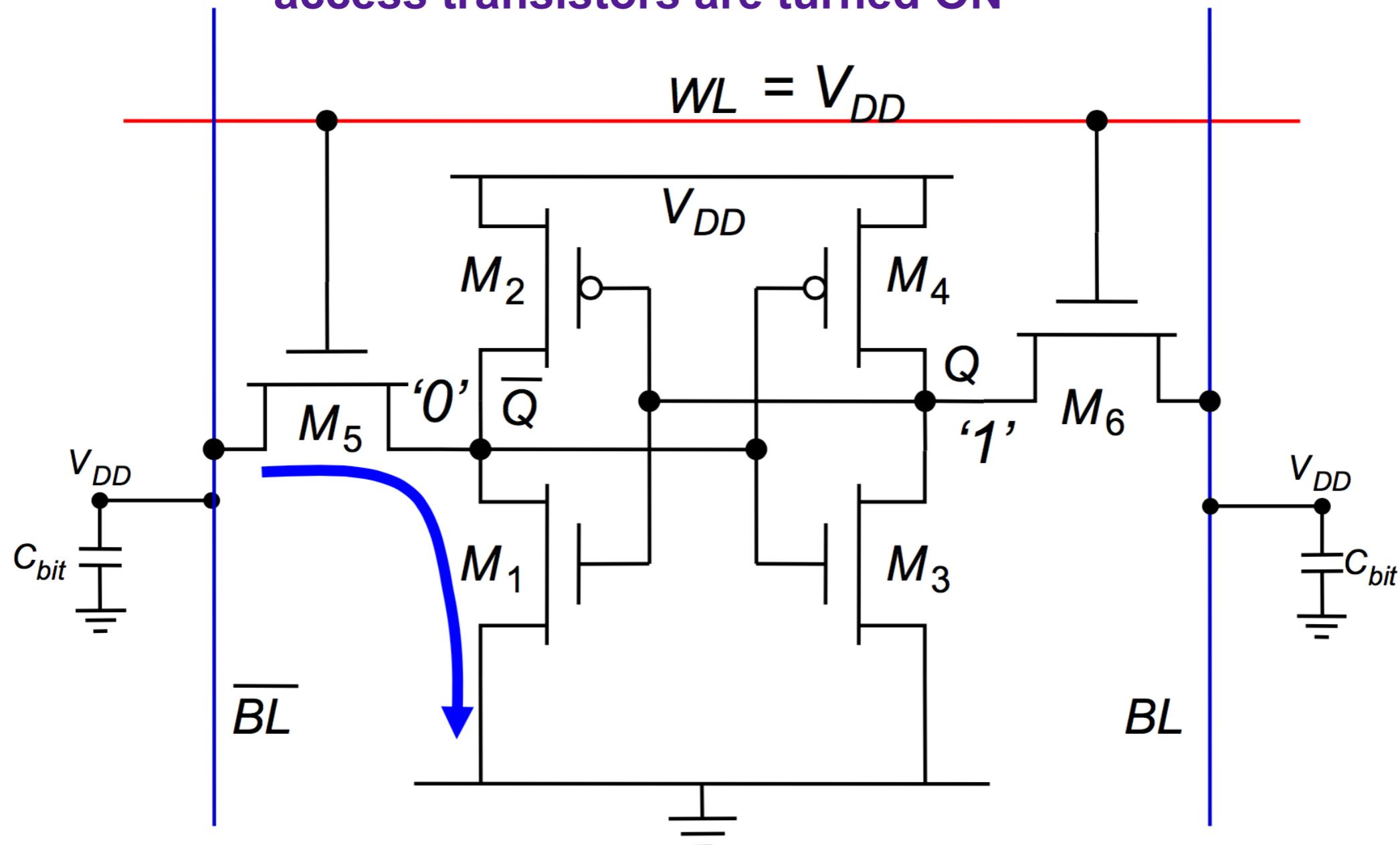
READ analysis



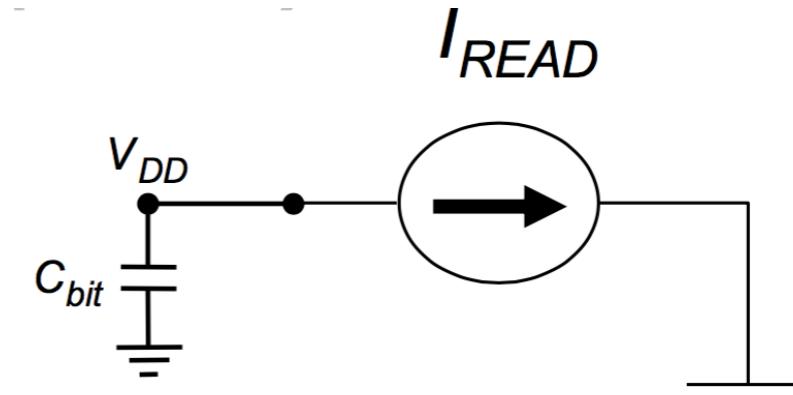
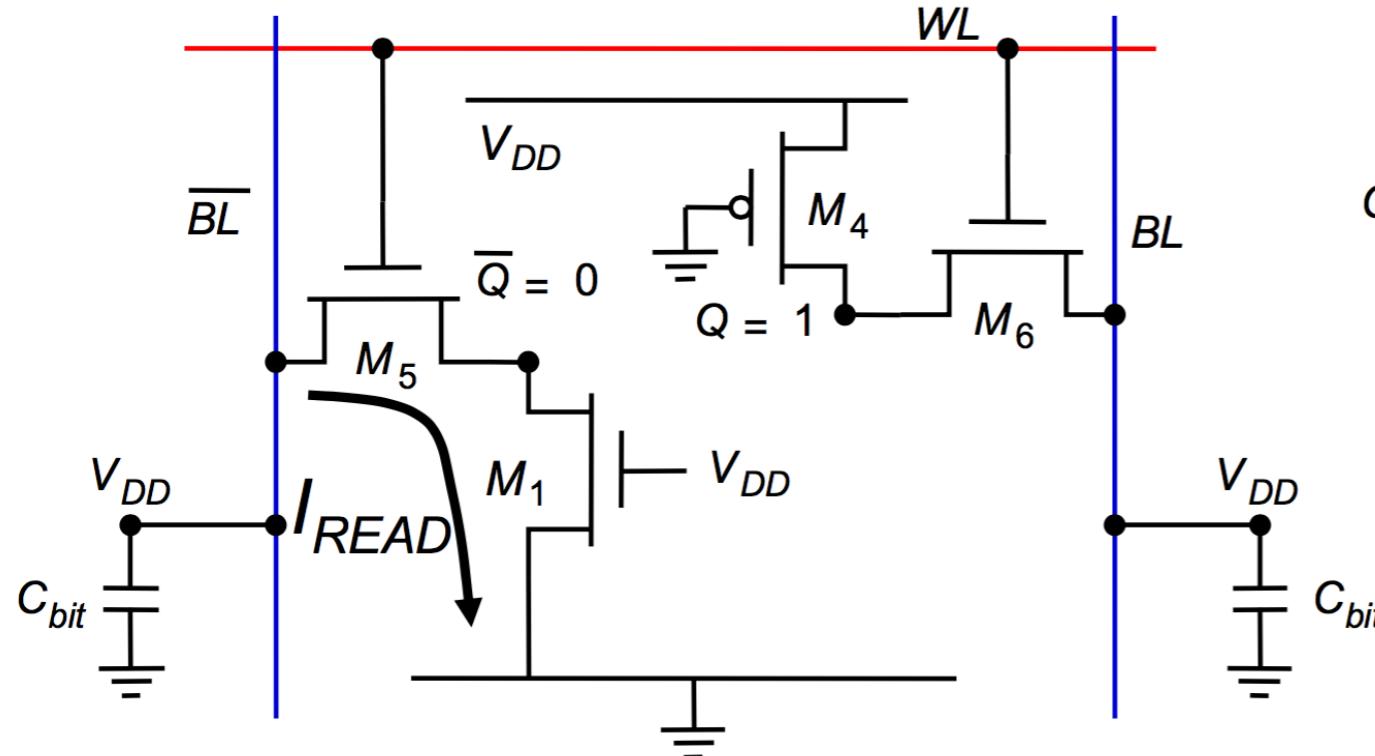
- During pre-charge, both bit line and its complement are charged to V_{DD} .

READ analysis

Read phase ($\text{PRE} = 1, \text{WL} = 1$) \rightarrow
access transistors are turned ON



READ analysis



$$C_{bit} \frac{dV_{bl}}{dt} + I_{READ} = 0$$

$$C_{bit} \frac{\Delta_{BIT}}{I_{READ}} = T_{access}$$

I_{READ} = read current

T_{access} = Access Time

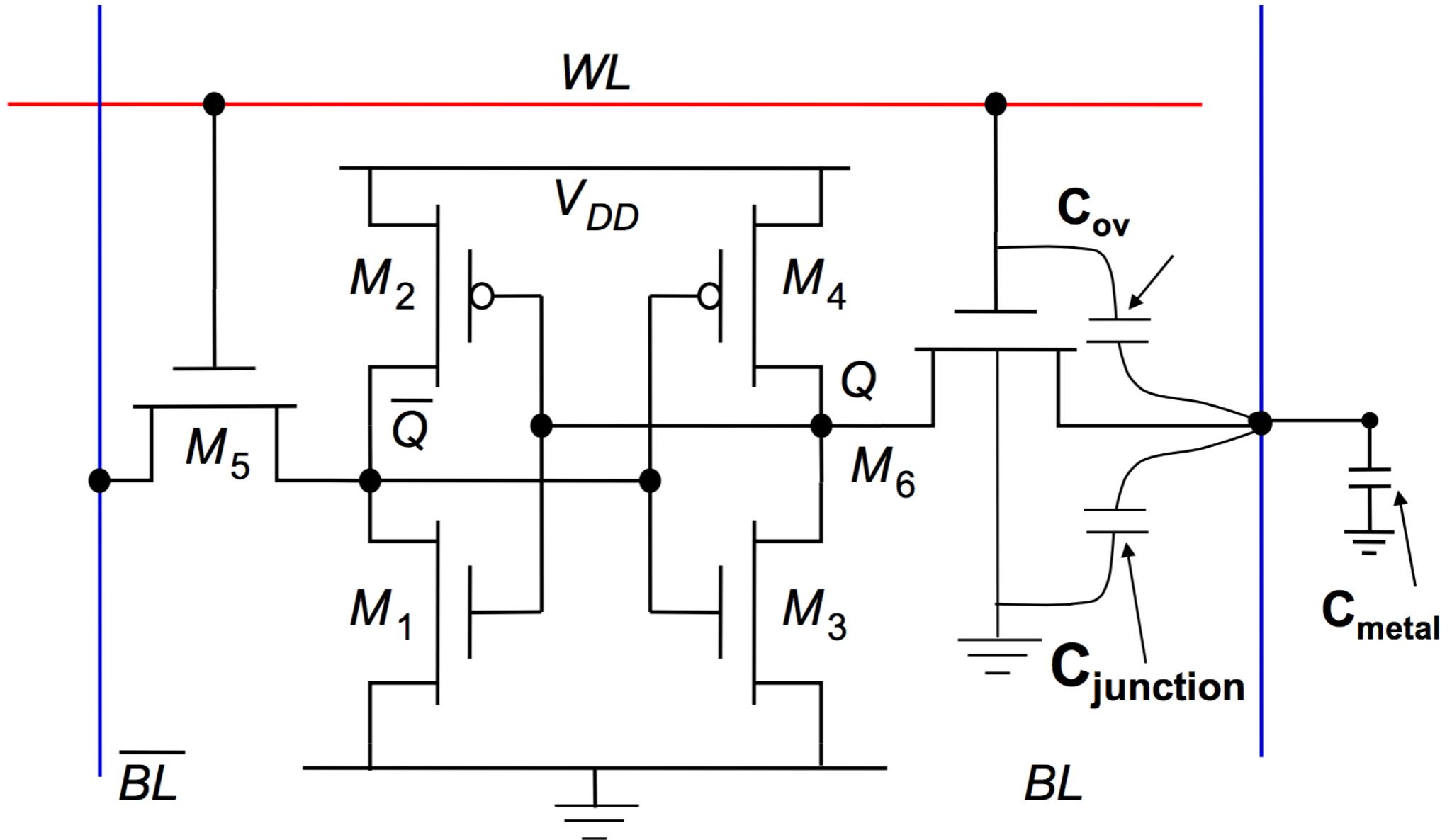
M_1 in linear region

Δ_{BIT} = bit - differential

Assume, bitline voltage drops slowly

=> M_5 is in saturation

Bitline capacitances



$$C_{bit_percell} = (C_{junction} + C_{ov} + C_{metal})$$

Bitline capacitances

- Junction capacitance of access transistors
- Gate-to-drain overlap capacitances
- Metal capacitance of the bit-line

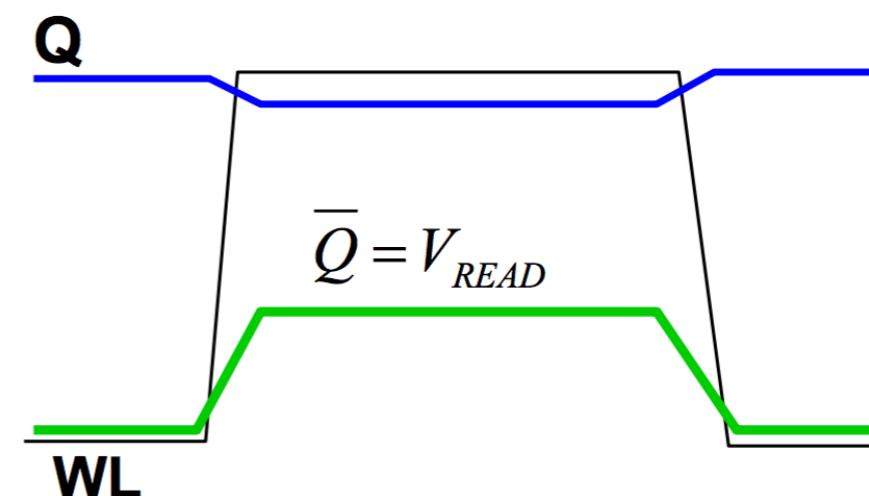
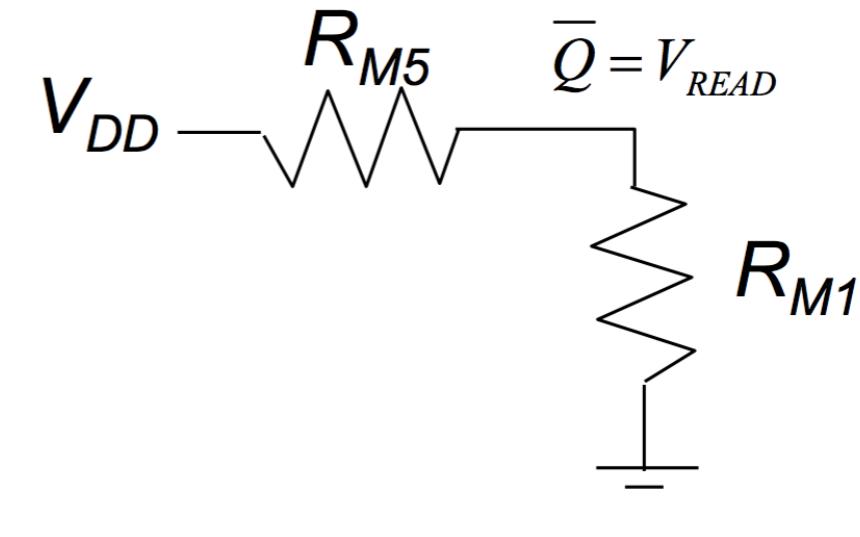
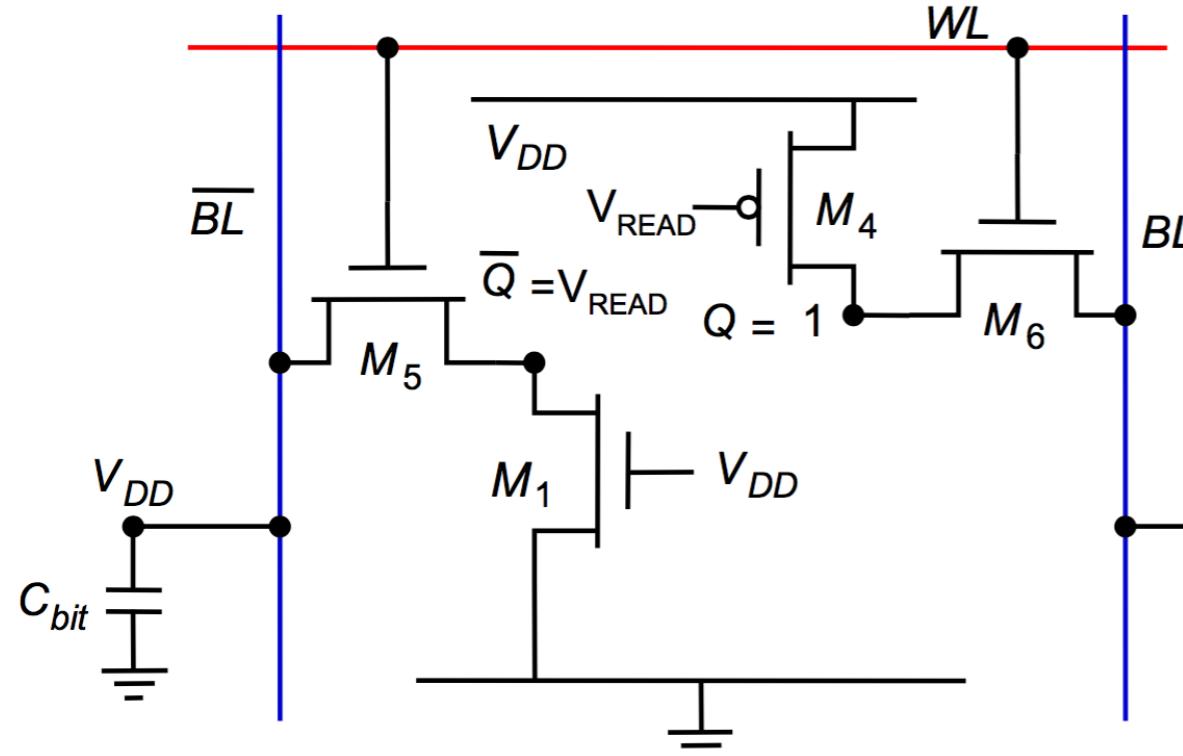
$$C_{bit} = N_{row} \left(C_{junction} + C_{ov} + C_{metal} \right)$$

$C_{junction}$ = junction cap of each access device

C_{ov} = gate – to – drain overlap cap of each access device

C_{metal} = metal cap of per cell

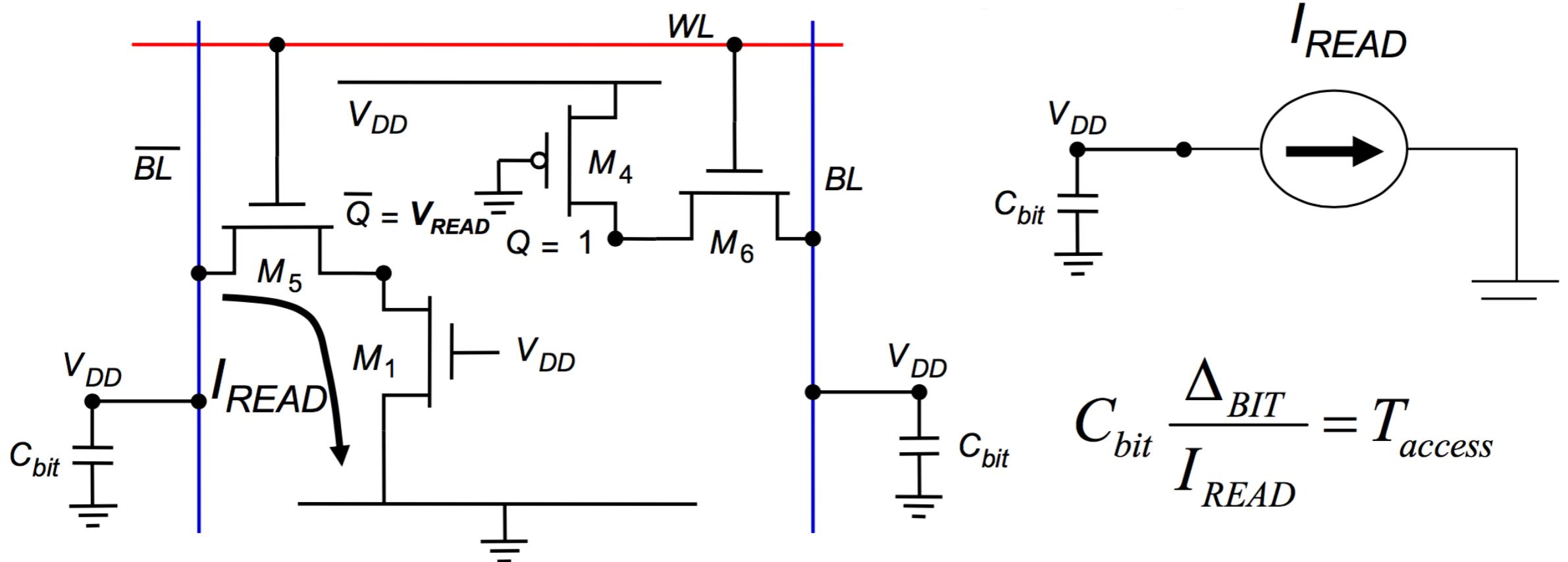
READ analysis



M_5 and M_1 from a resistor divider

V_{READ} =read voltage

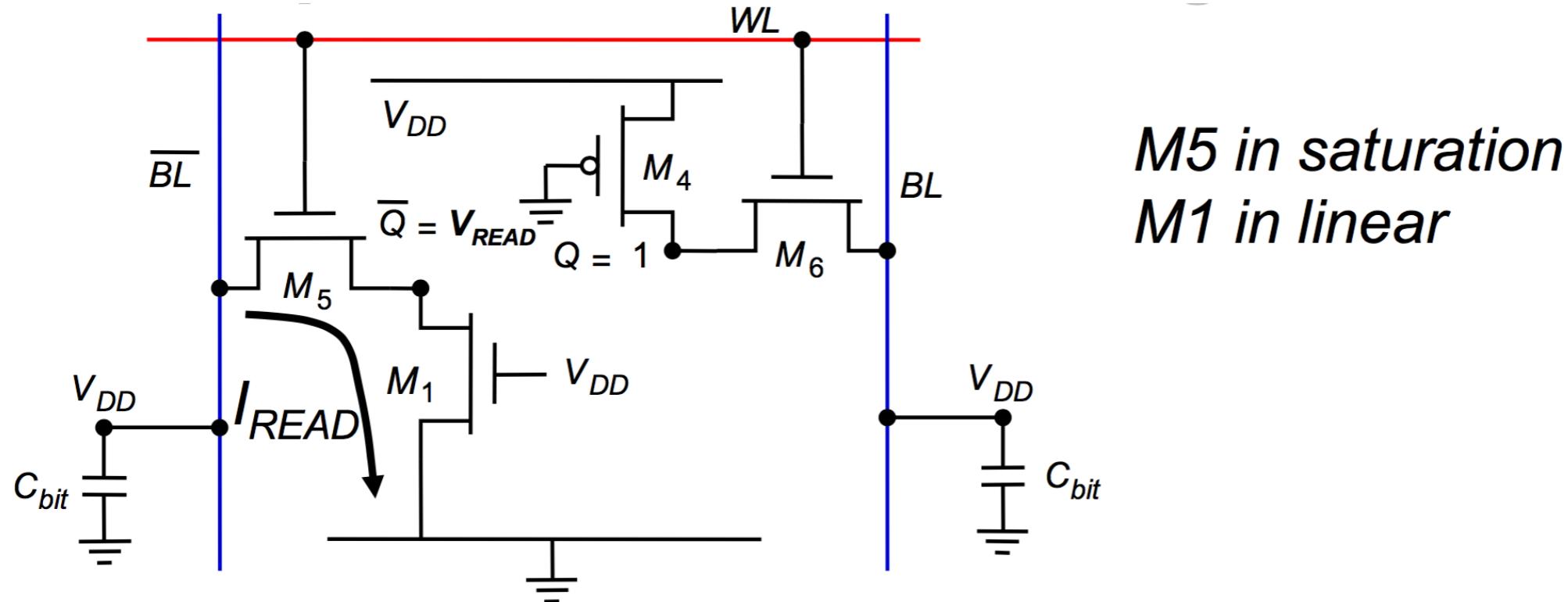
READ analysis



$$I_{READ} = \mu_n C_{ox} \frac{W_{M5}}{2L_{M5}} (V_{DD} - V_{READ} - V_{th})^2 = 0.5 \beta_{M5} (V_{DD} - V_{READ} - V_{th})^2$$

$$T_{ACCESS} = C_{bit} \frac{\Delta_{BIT}}{0.5 \beta_{M5} (V_{DD} - V_{READ} - V_{th})^2}$$

Calculation of read voltage

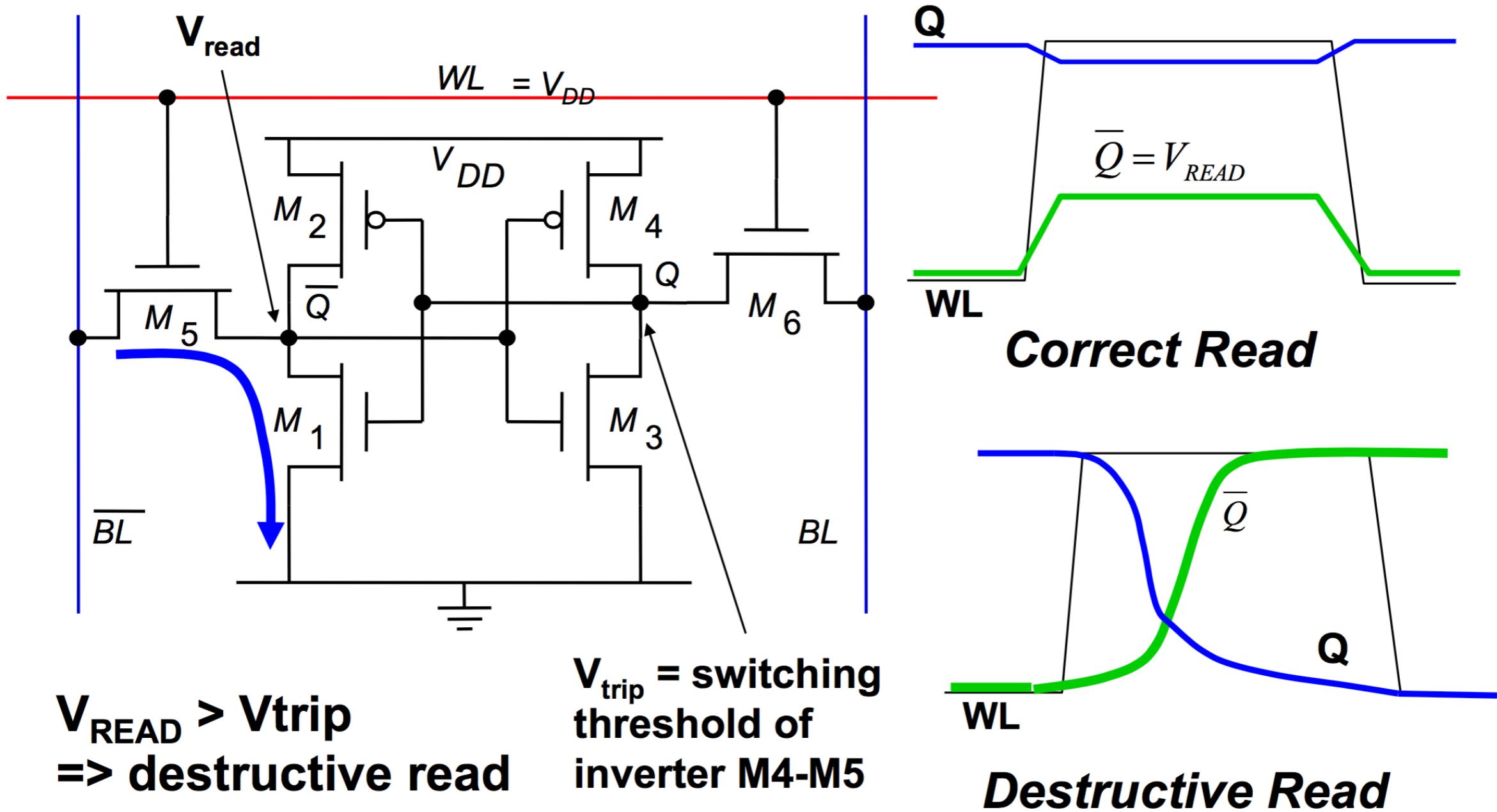


$$0.5\beta_{M5} (V_{DD} - V_{READ} - V_{th})^2 = \beta_{M1} [(V_{DD} - V_{th}) - 0.5V_{READ}] V_{READ}$$

$$\left(1 + \underbrace{\beta_{M1}/\beta_{M5}}_{\beta_{ratio-pd-ax}} \right) V_{READ}^2 - 2(1 + \beta_{M1}/\beta_{M5})(V_{DD} - V_{th})V_{READ} + (V_{DD} - V_{th})^2 = 0$$

$$V_{READ} = (V_{DD} - V_{th}) \left(1 - \sqrt{1 - 1/(1 + \beta_{ratio-pd-ax})} \right)$$

Destructive read



Noise margin for read

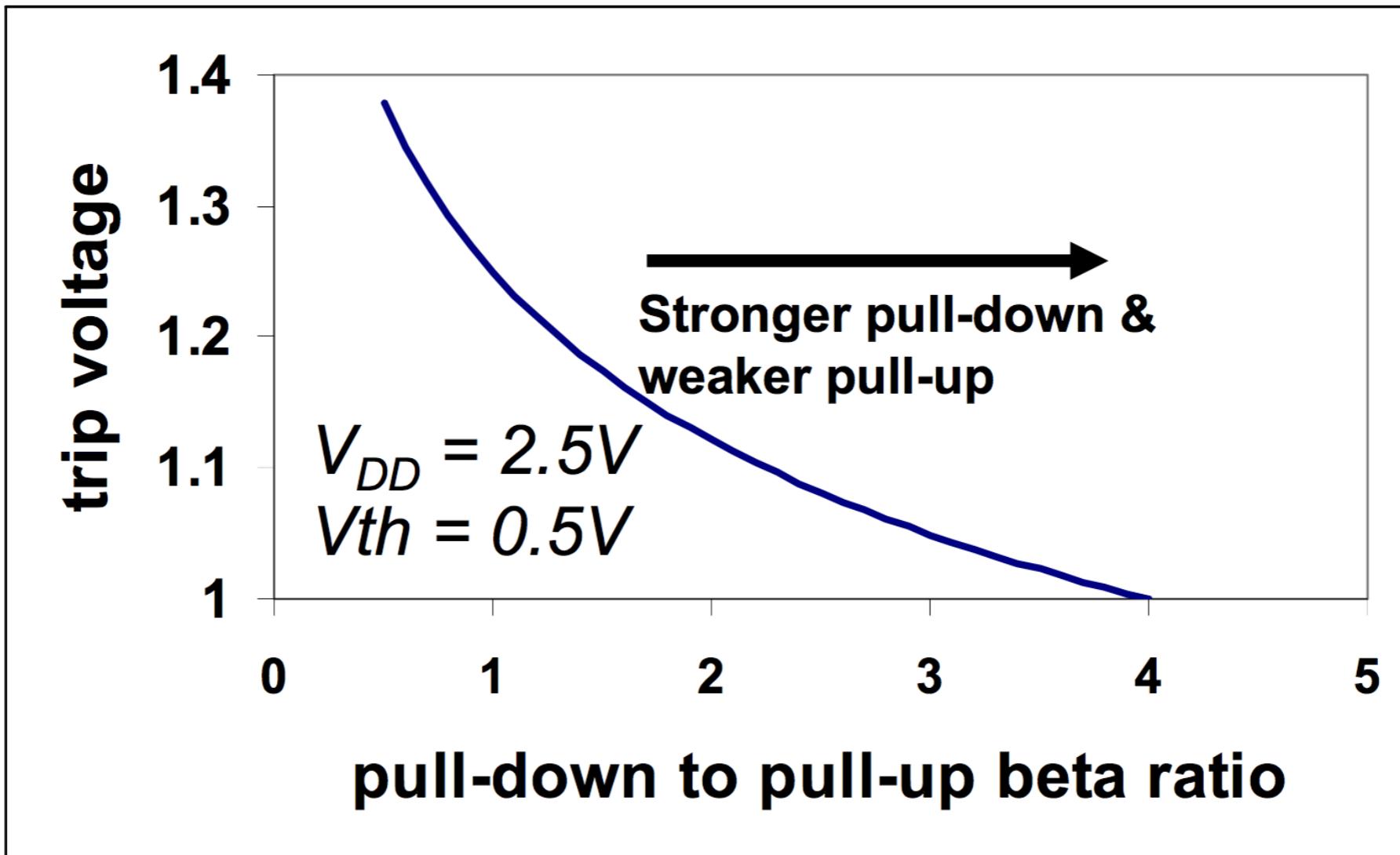
A simple definition (can be used for first-order design)

$$\text{Read Margin} = V_{trip} - V_{read}$$

$$V_{trip} = \frac{V_{DD} - |V_{thp}| + \sqrt{\beta_{M3}/\beta_{M4}} V_{thn}}{1 + \sqrt{\beta_{M3}/\beta_{M4}}} \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{trip} \\ \text{stronger pull-up} \Rightarrow \text{higher } V_{trip} \end{array} \right\}$$
$$= \frac{V_{DD} - |V_{thp}| + \sqrt{\beta_{ratio-pd-pup}} V_{thn}}{1 + \sqrt{\beta_{ratio-pd-pup}}} \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{trip} \\ \text{stronger pull-up} \Rightarrow \text{higher } V_{trip} \end{array} \right\}$$

$$V_{READ} = (V_{DD} - V_{thn}) \left(1 - \sqrt{1 - \frac{1}{1 + \beta_{ratio-pd-ax}}} \right) \quad \left. \begin{array}{l} \text{stronger pull-down} \Rightarrow \text{lower } V_{read} \\ \text{stronger access} \Rightarrow \text{higher } V_{read} \end{array} \right\}$$

Trip voltage



- A stronger pull-up device increases V_{trip} .
- A weaker pull-down device increases V_{trip} .

Read access speed

For lower read access time:

Higher I_{read}

- ⇒ stronger access & stronger pull-down
- ⇒ Pull-up device does not contribute to read speed
- ⇒ Pull-up can be made weaker (i.e. of smaller width) which will also reduce area

Read noise margin

For better read margin:

Smaller V_{read} => Higher $\beta_{ratio-pd-ax}$

=> weaker access & stronger pull-down

Higher V_{trip} => Lower $\beta_{ratio-pd-pup}$

=> weaker pull-down & stronger pull-up

- **Read margin requirement for access device contradicts with read speed requirement**
- **Read voltage requirement for pull-down device contradicts trip point requirement**
 - Read voltage is more sensitive to pull-down strength compared to trip voltage
 - Generally a larger pull-down helps read margin but it cannot be too large

Class on Oct. 26, 2015

- Project discussion
 - choose team members, groups of 4.
- Recap of materials for exam
- Homework # 4 will be due on Oct. 26, 2015