CS 152
Computer Architecture and Engineering

Lecture 12 – Memory and Interfaces

2005-10-11

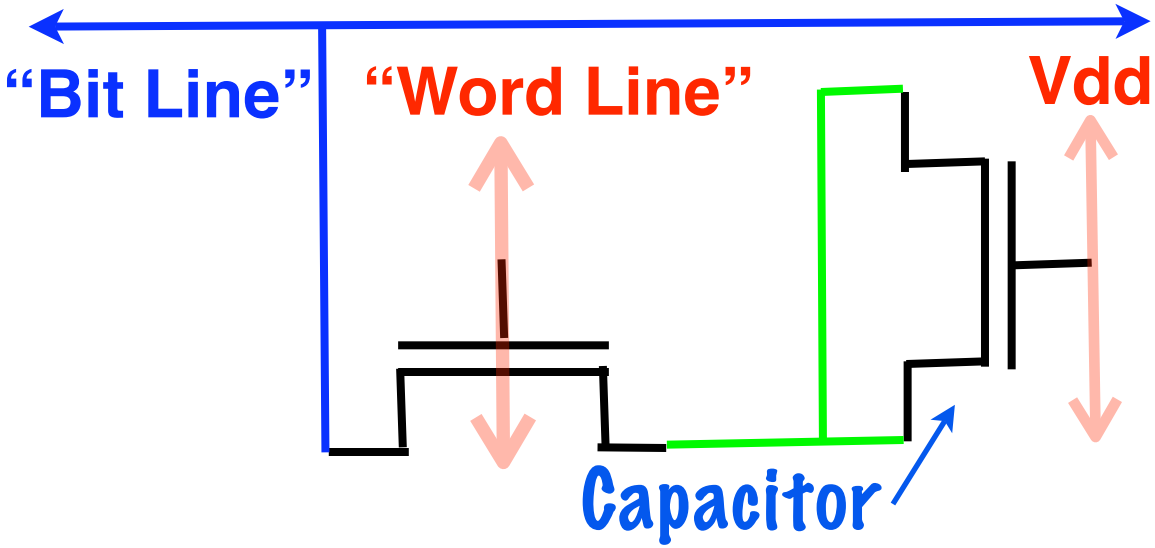
John Lazzaro
(www.cs.berkeley.edu/~lazzaro)

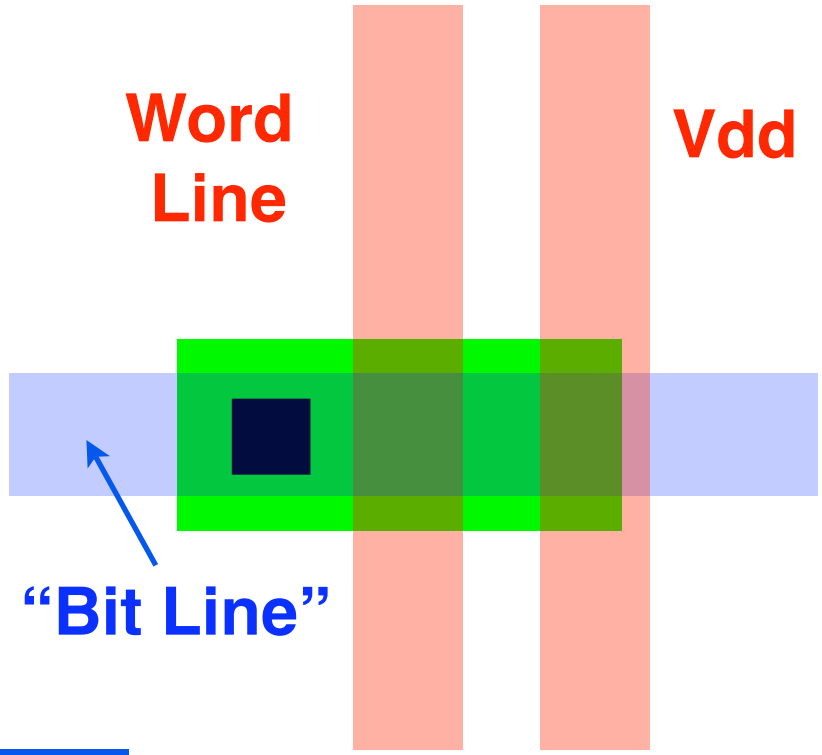
TAs: David Marquardt and Udam Saini

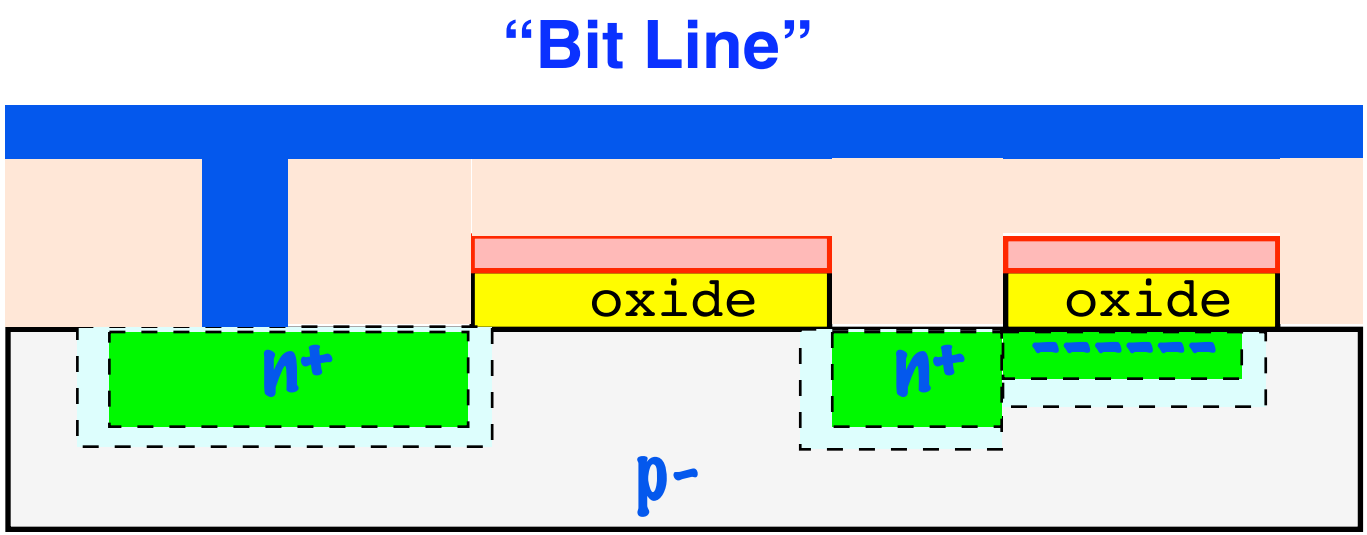
www-inst.eecs.berkeley.edu/~cs152/

UC Regents Fall 2005 © UCB

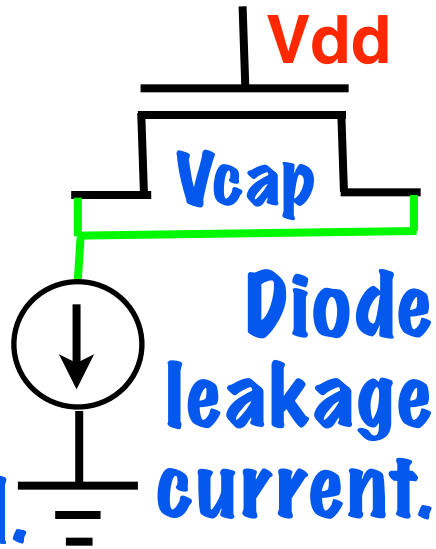
Last Time: 1-T DRAM cells







Why Vcap values start out at ground.



Diode leakage current.

Word Line and Vdd run on "z-axis"

UC Regents Fall 2005 © UCB

Today: Memory Technology Wrap-Up

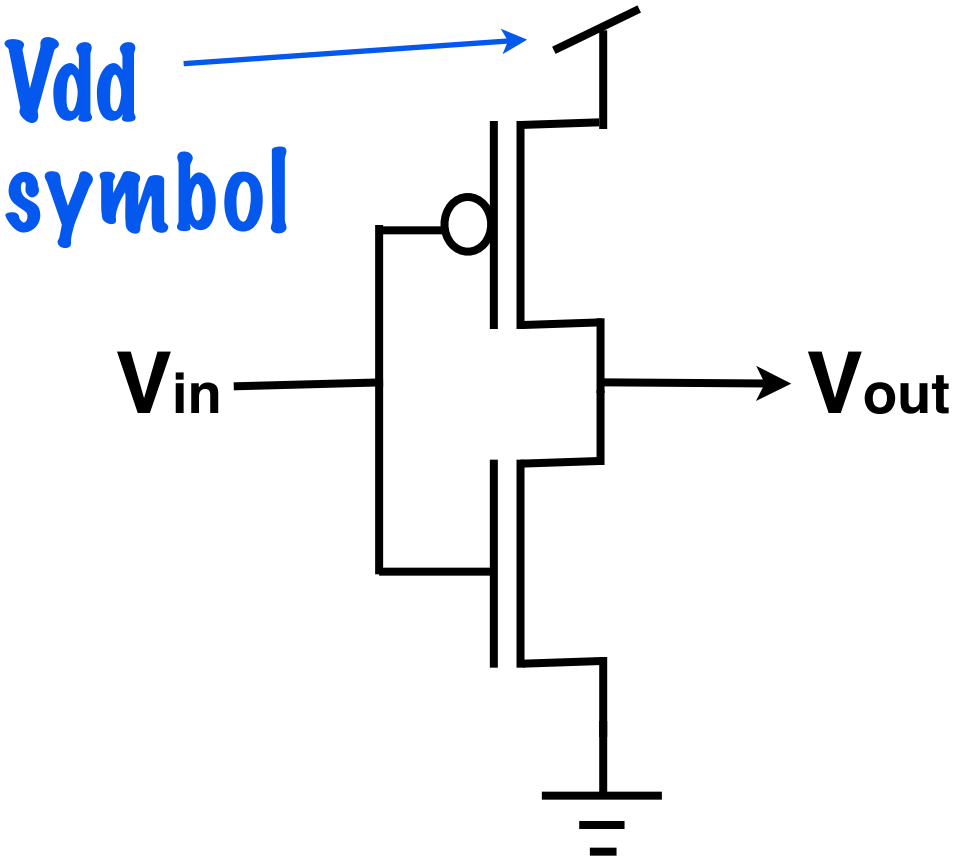
- **Static Memory Circuits:** For SRAM memory cells and for flip-flops.
- **Memory Arrays:** Row decoders, column sense amps, array sizing.
- **DRAM Interfaces:** How the SDRAM chips on the Calinx board work.

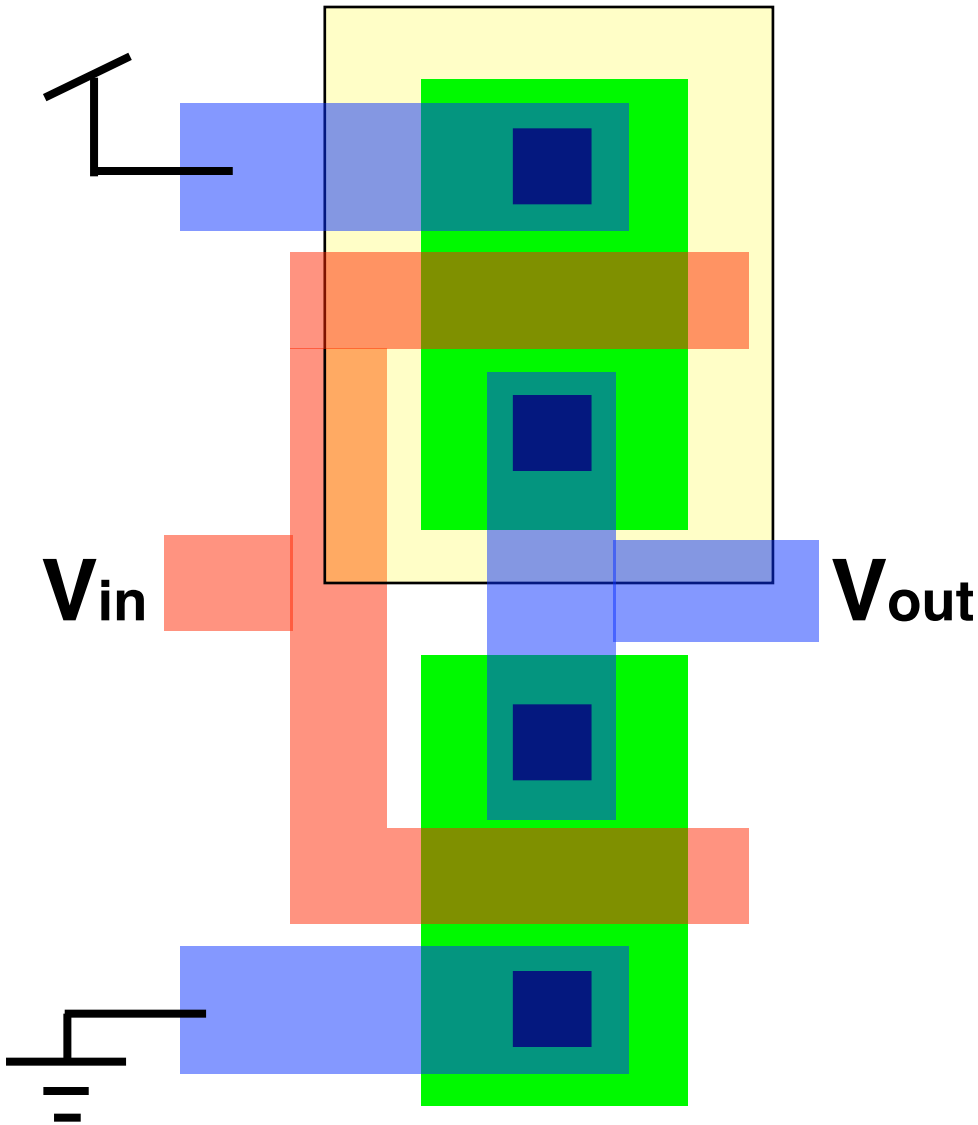
UC Regents Fall 2005 © UCB

Inverters

UC Regents Fall 2005 © UCB

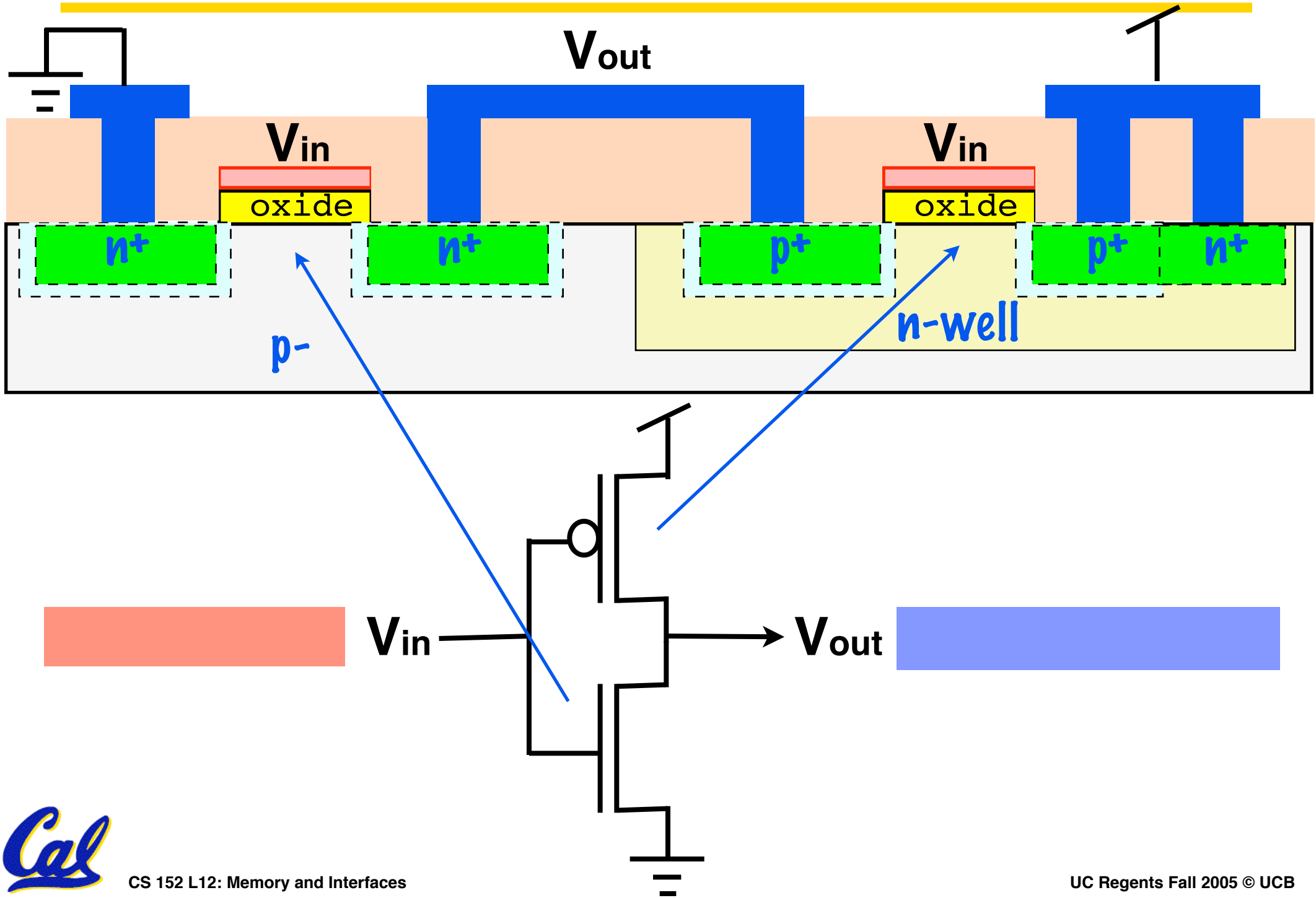
Inverters: Circuits and Layout





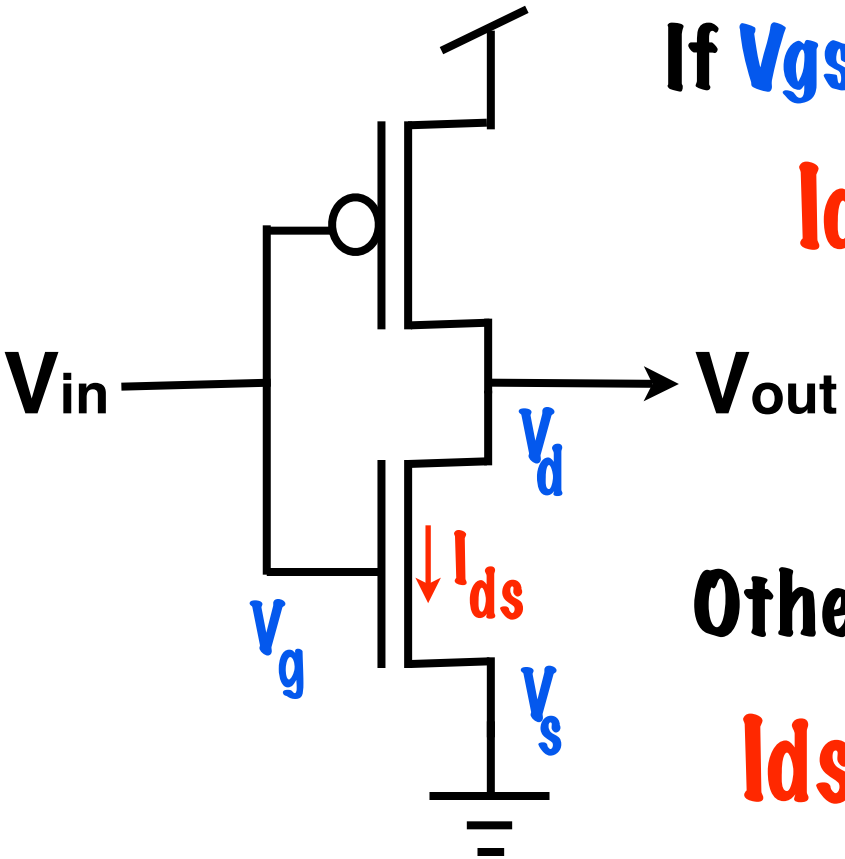
UC Regents Fall 2005 © UCB

Inverter: Die Cross Section



UC Regents Fall 2005 © UCB

Inverters: n-fet Transistor Equation



If Vgs > Vt and Vds > Vgs - Vt :

$$I_{ds} = (k/2)\,(W/L)\,[V_{gs} - V_t]^2$$

Otherwise, if Vgs > Vt :

$$I_{ds} = k\,(W/L)\,[V_{gs} - V_t]\,[V_{ds}]$$

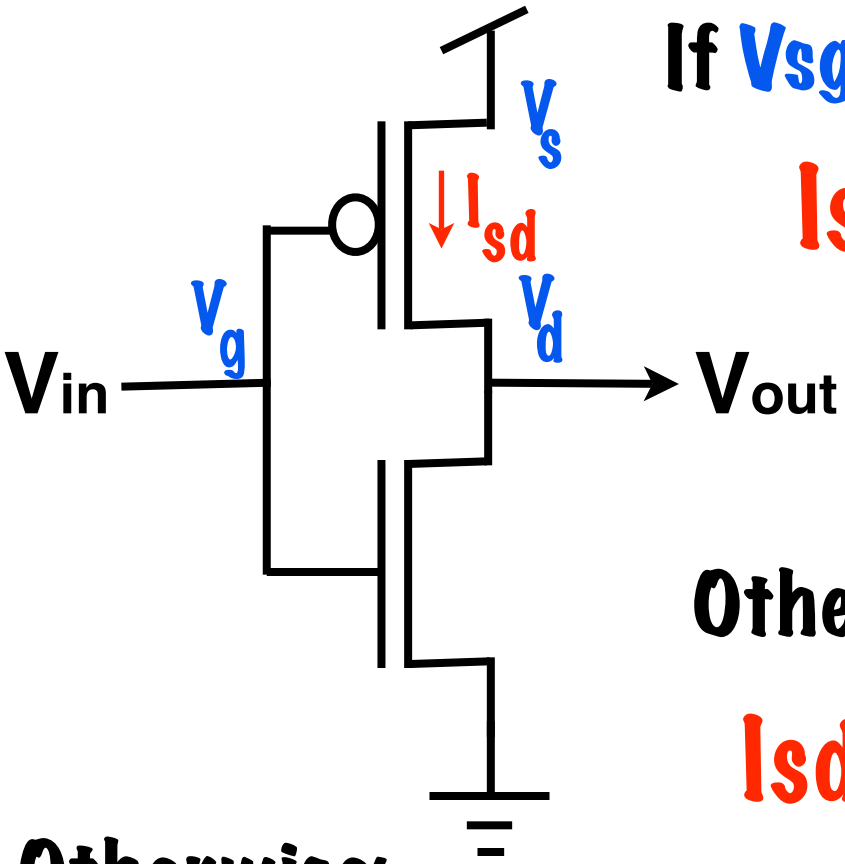
Otherwise:

$$I_{ds} \approx 0, \text{ but really} = I_0\,[\exp((\kappa V_g - V_s)/V_0)]\,[1 - \exp(-V_{ds}/V_0)]$$

Note: Vt is transistor threshold, was formerly Vth. Also, Vt is actually Vt(Vs) ~sqrt(Vs).

UC Regents Fall 2005 © UCB

Inverters: p-fet Transistor Equation



If Vsg > Vt and Vsd > Vsg - Vt :

$$Isd = (k/2)\,(W/L)\,[Vsg - Vt]^2$$

Otherwise, if Vsg > Vt :

$$Isd = k\,(W/L)\,[Vsg - Vt]\,[Vsd]$$

Otherwise:

Isd ≈0, but again, in reality there is a "leakage" current.

Note: Vt for p-Fet and n-Fet are different. Also true for "k" (fab constant). kp < kn, due to electrons being faster than holes.

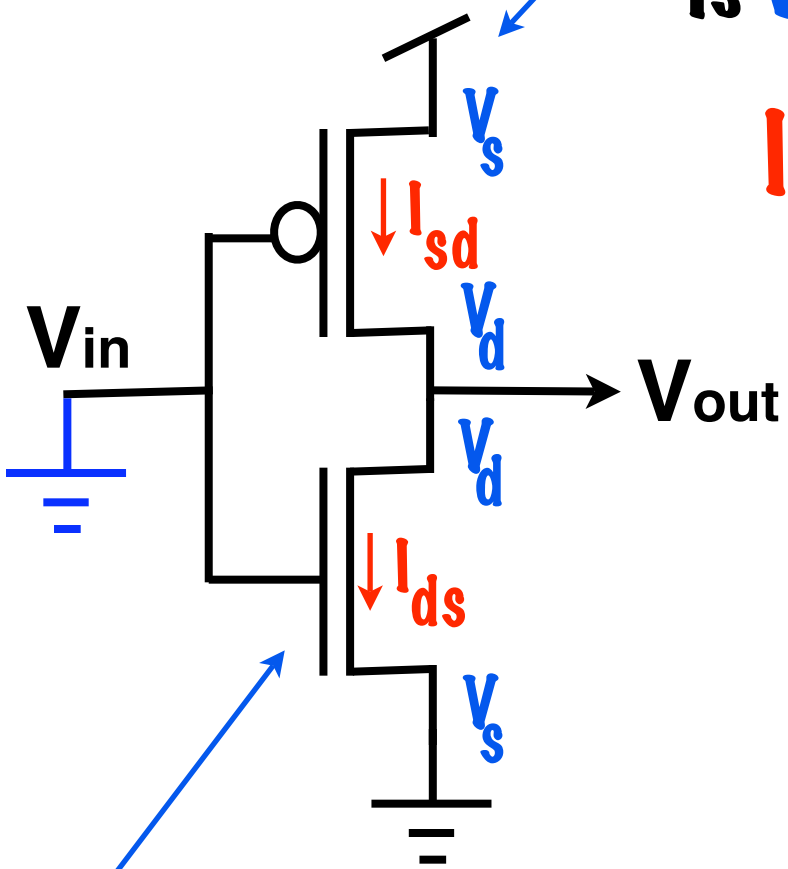
UC Regents Fall 2005 © UCB

Inverters with Vin = Gnd, Vout = Vdd

Is Vsd > Vsg - Vt once Vout is Vdd?

Is Vsg > Vt ?

$$I_{sd} = k\,(W/L)\,[V_{sg} - V_t]\,[V_{sd}]$$



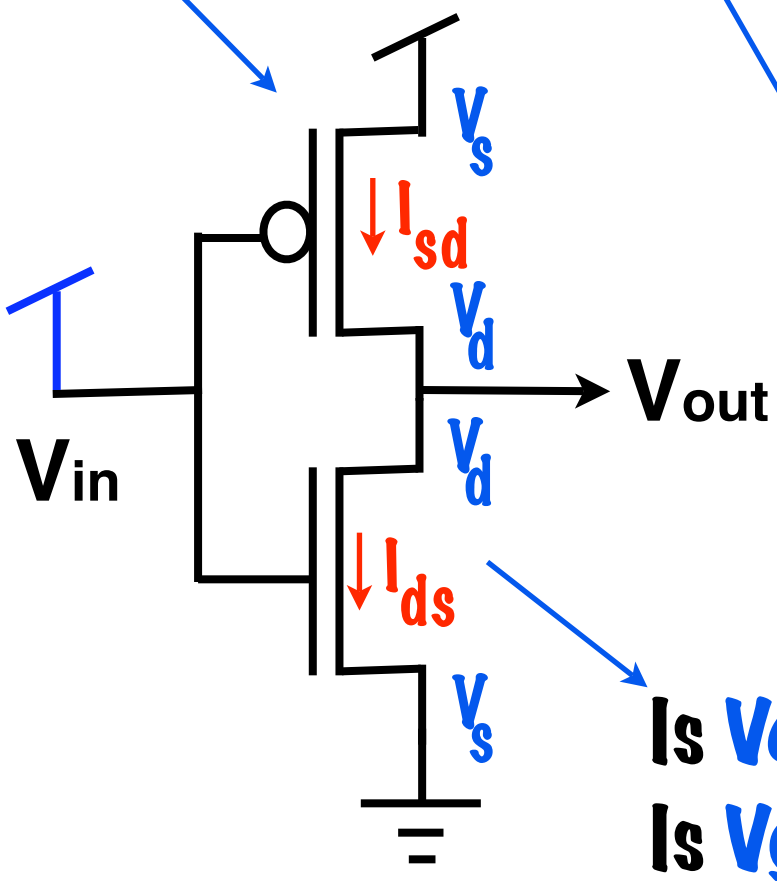
This goes as close to 0 as it can while still supplying the leakage current.

Ids ≈0, but really a small leakage current

UC Regents Fall 2005 © UCB

Inverters with Vin = Vdd, Vout = Gnd

Isd ≈0, but really a small leakage current



This goes as close to 0 as it can while still supplying the leakage current.

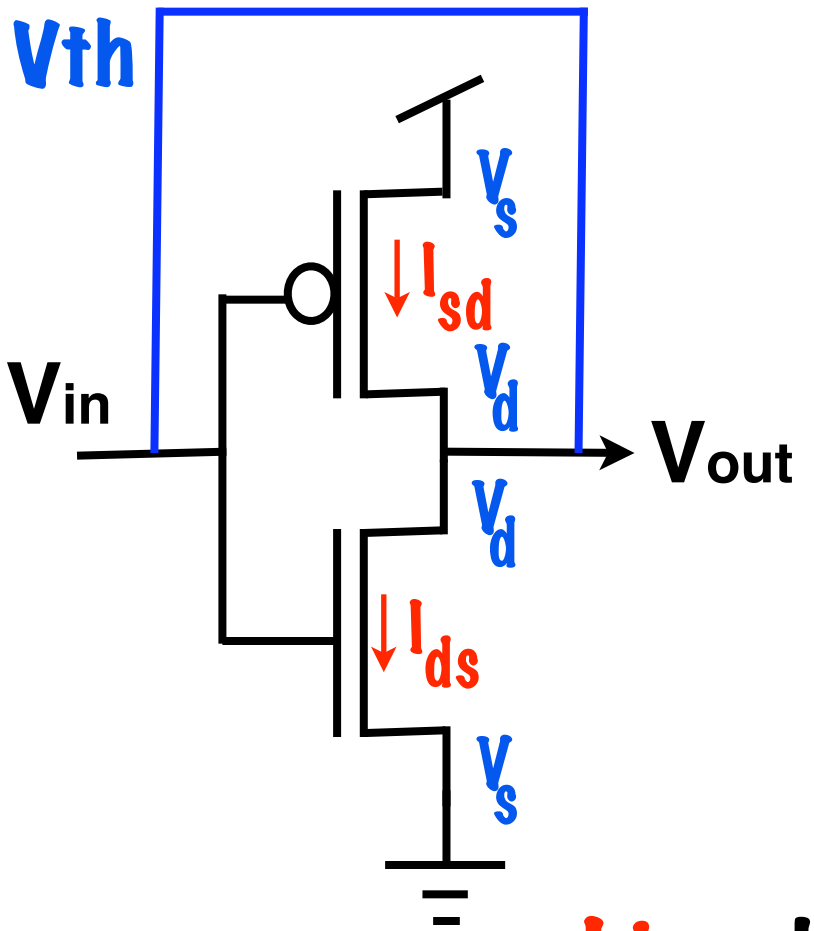
Is Vds > Vgs - Vt once Vout is Gnd?

Is Vgs > Vt ?

$$I_{ds} = k\,(W/L)\,[V_{gs} - V_t]\,[V_{ds}]$$

UC Regents Fall 2005 © UCB

Calculating the inverter threshold (Vth)



Tie output to input. Vth

Assume voltage is "somewhere near the middle"

For nfet, is Vds > Vgs - Vt ?

For pfet, is Vsd > Vsg - Vt ?

No, by definition! Use:

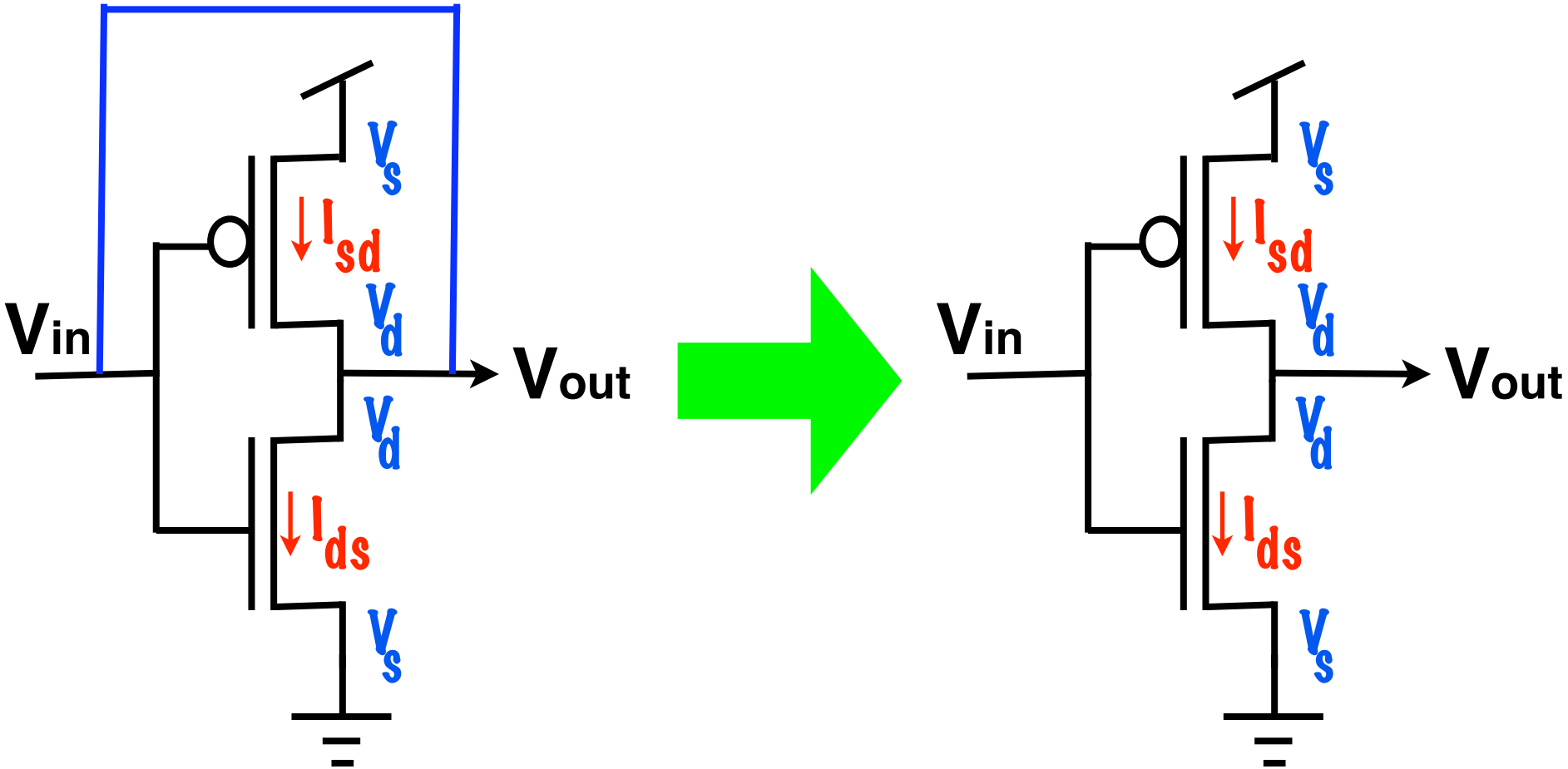
$$Ids = kn\,(W/L)\,[Vth - Vtn]\,[Vth]$$

$$Isd = kp\,(W/L)\,[Vdd - Vth - Vtp]\,[Vdd - Vth]$$

To compute the exact "voltage in the middle".

UC Regents Fall 2005 © UCB

Question: What happens when ...



Stays at Vth until a tiny amount of Vin noise appears.

Then output goes to Vdd or Gnd until ...

... Vin noise flips it back the other way.

Lesson: at Vth, small dVin make big dVout

UC Regents Fall 2005 © UCB

Static Memory Circuits

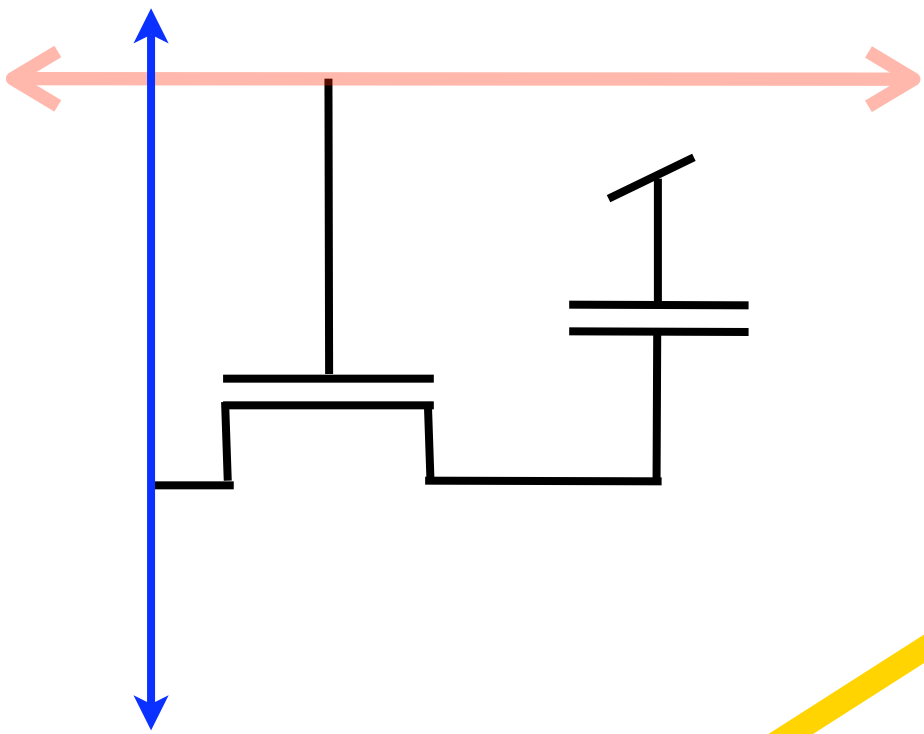
Dynamic Memory: Circuit remembers for a fraction of a second.

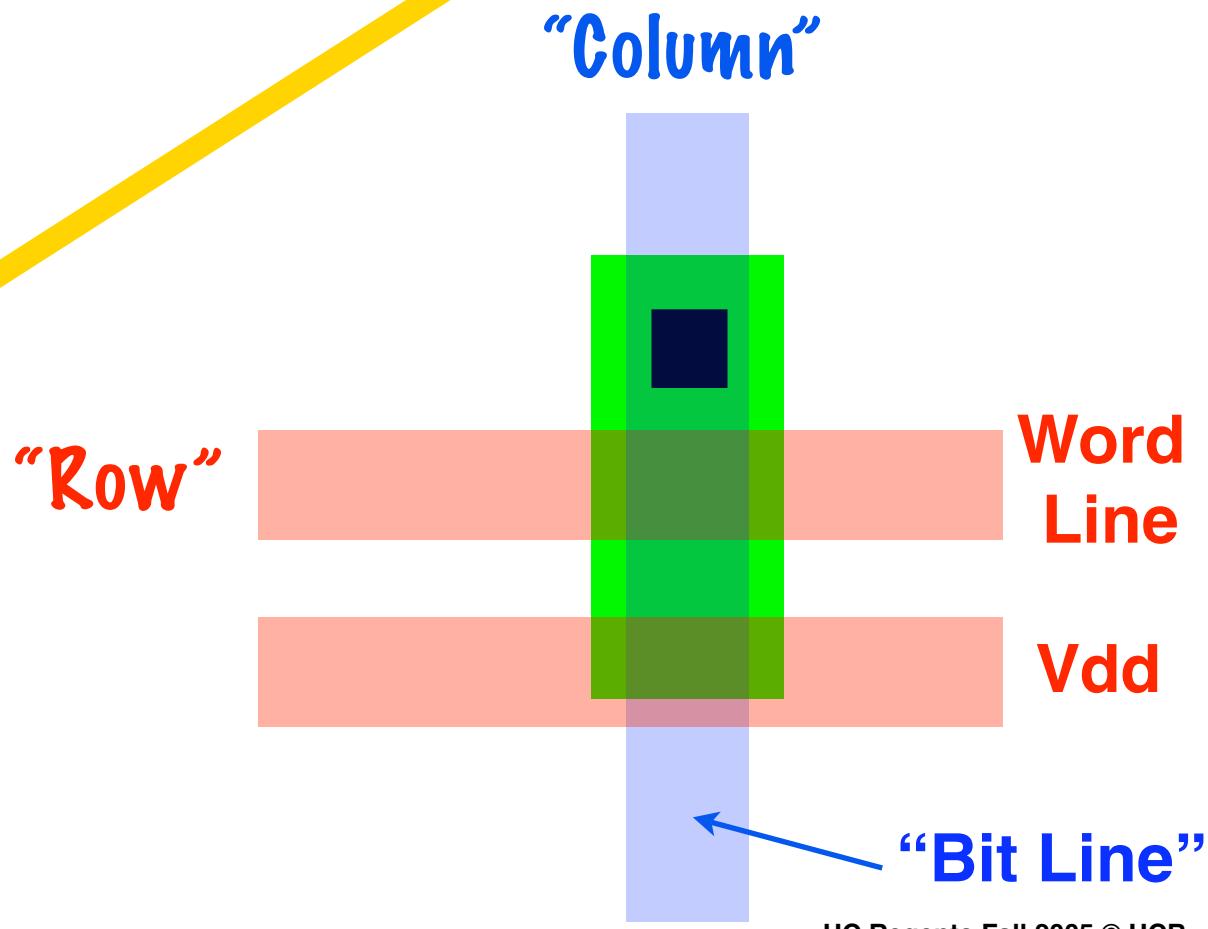
Static Memory: Circuit remembers as long as the power is on.

Non-volatile Memory: Circuit remembers for many years, even if power is off.

UC Regents Fall 2005 © UCB

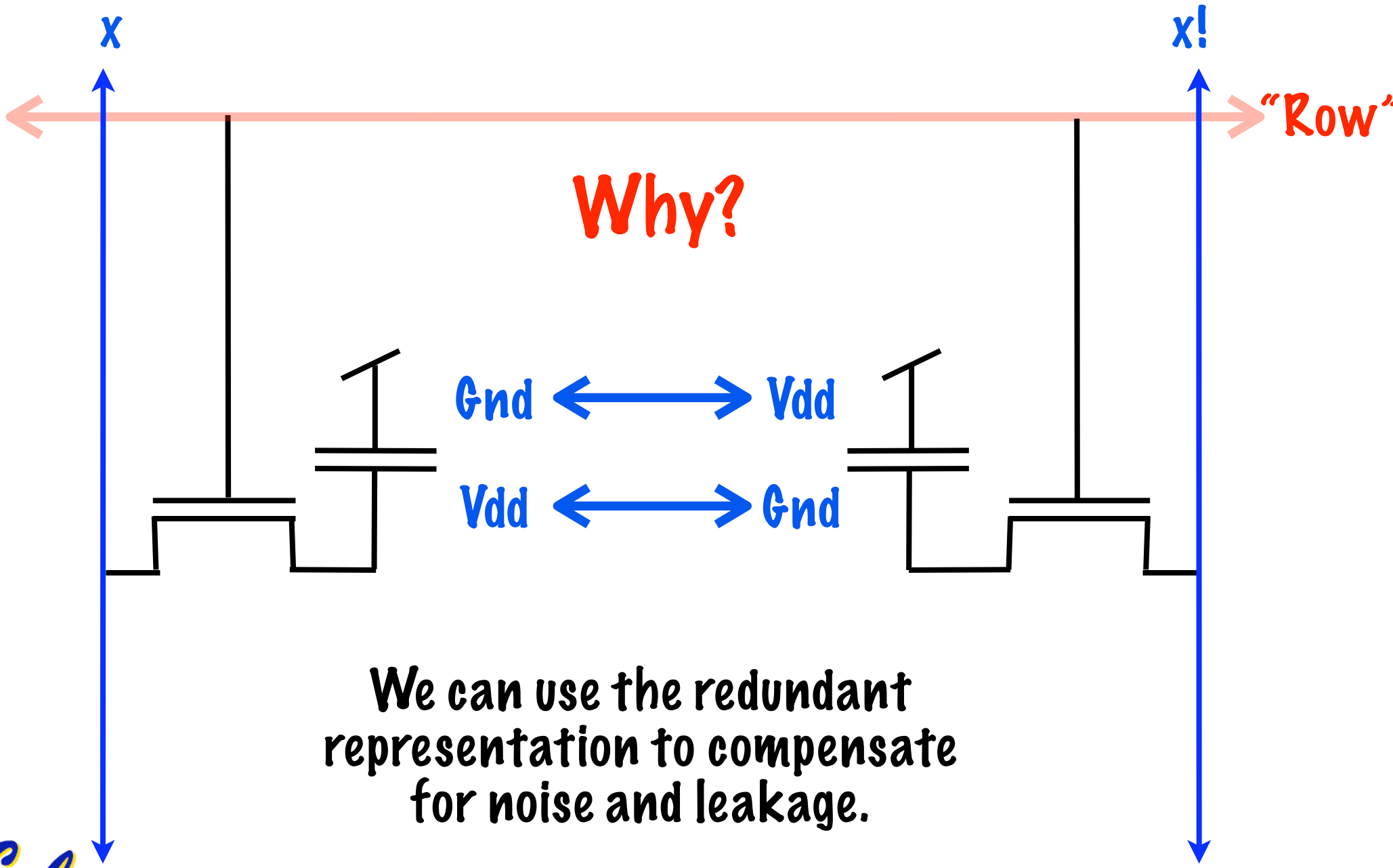
Recall DRAM cell: 1 T + 1 C





UC Regents Fall 2005 © UCB

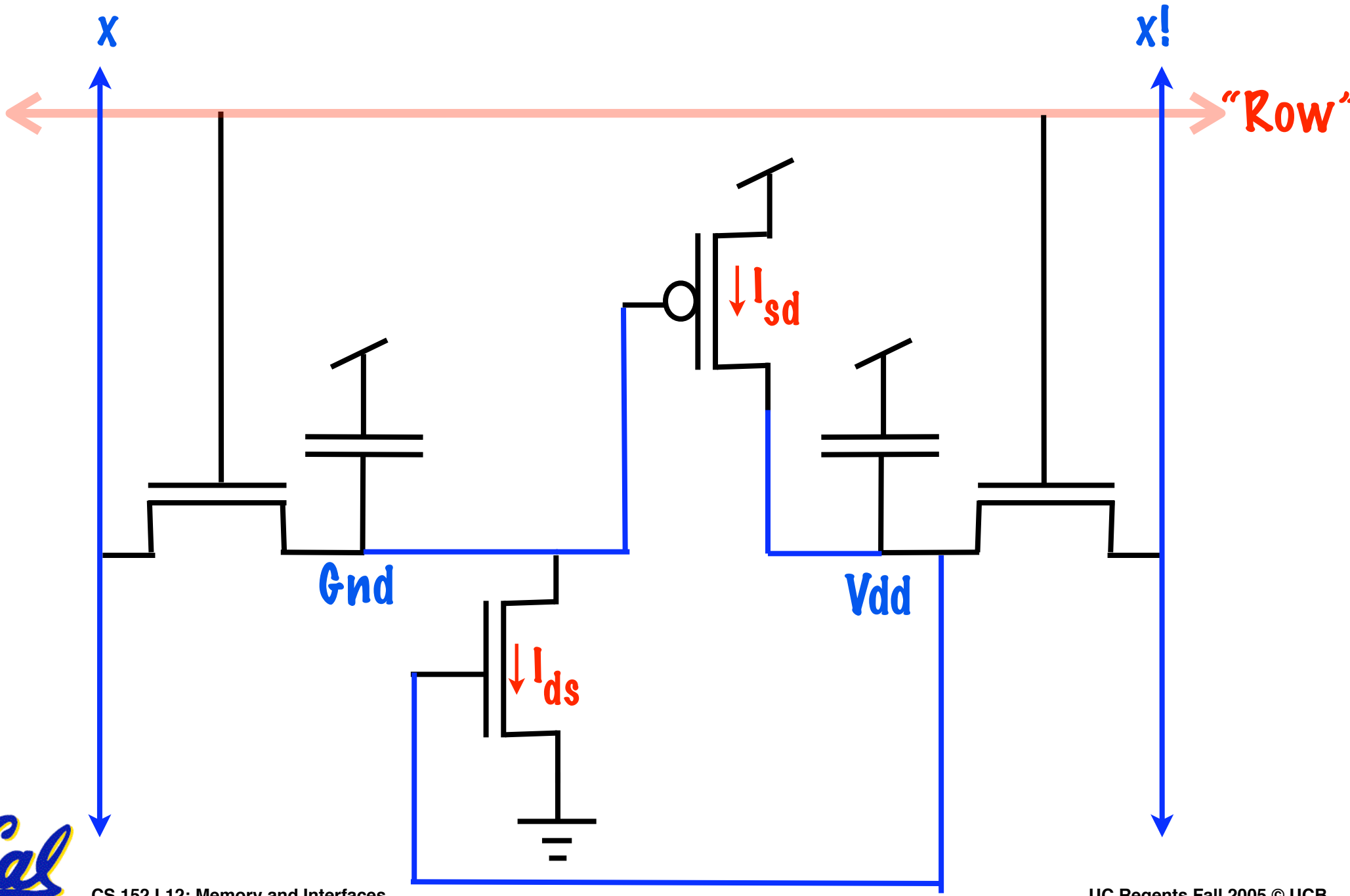
Idea: Store each bit with its complement



We can use the redundant representation to compensate for noise and leakage.

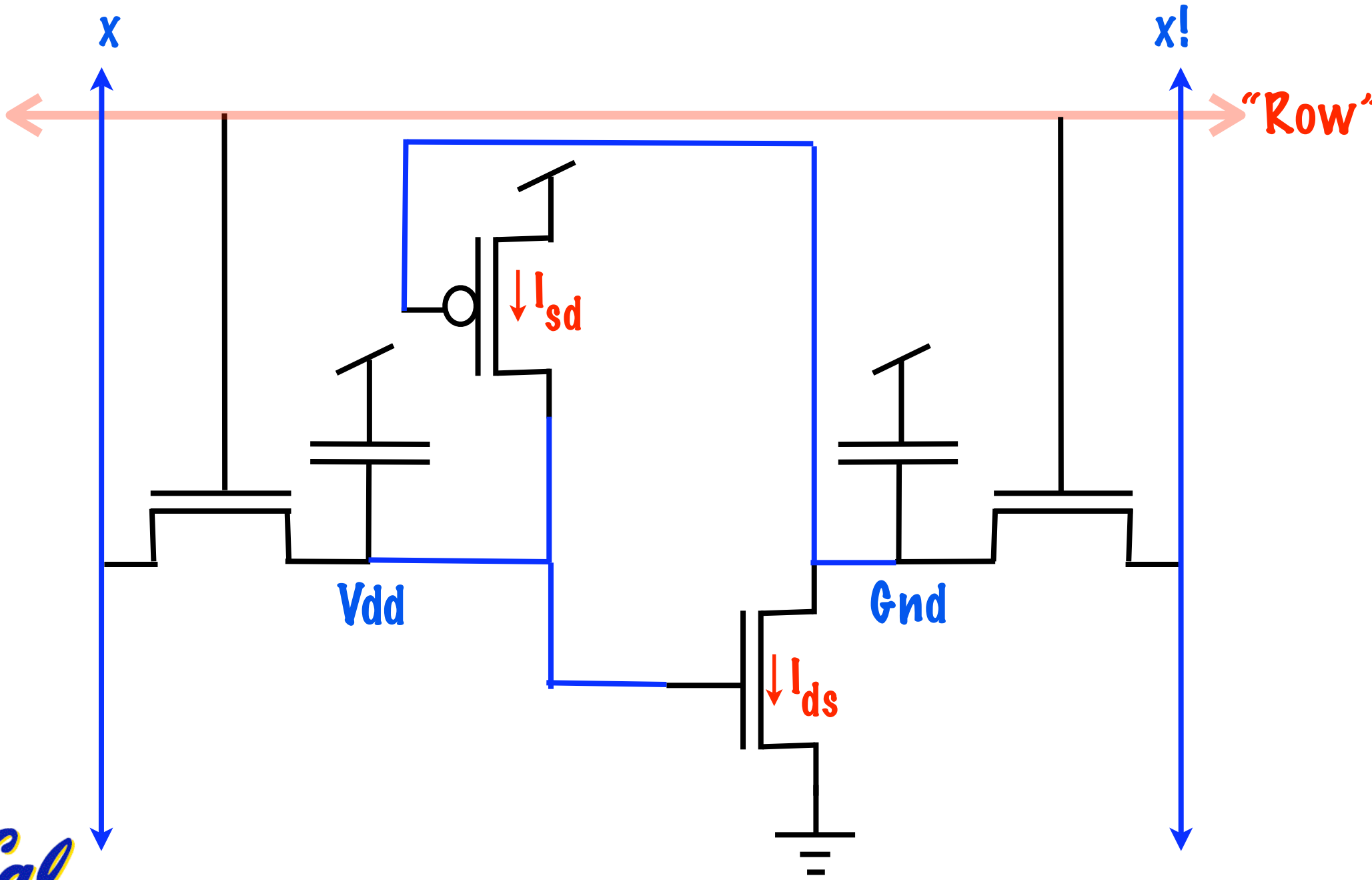
UC Regents Fall 2005 © UCB

Case #1: x = Gnd, x! = Vdd ...



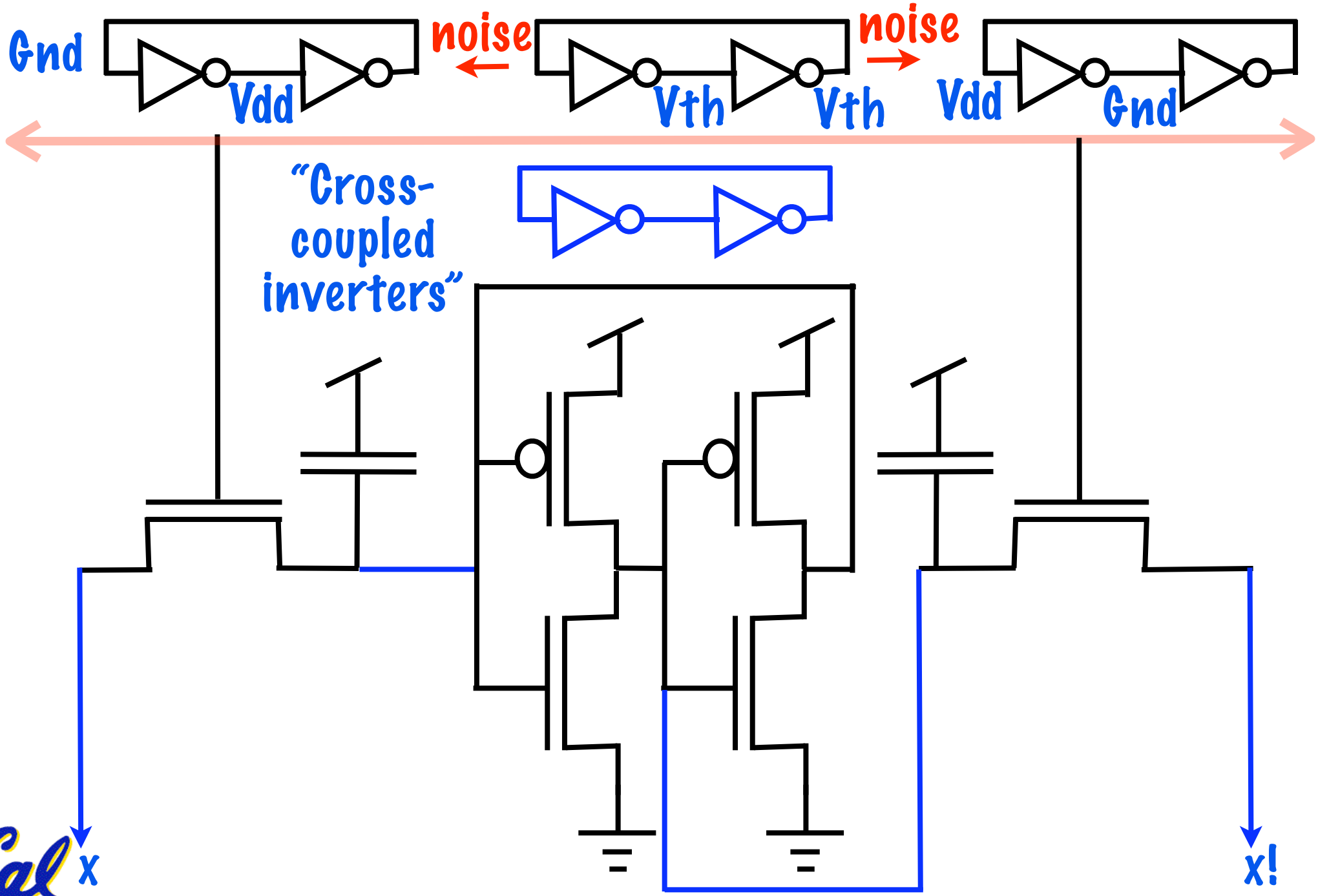
UC Regents Fall 2005 © UCB

Case #2: x = Vdd, x! = Gnd ...



UC Regents Fall 2005 © UCB

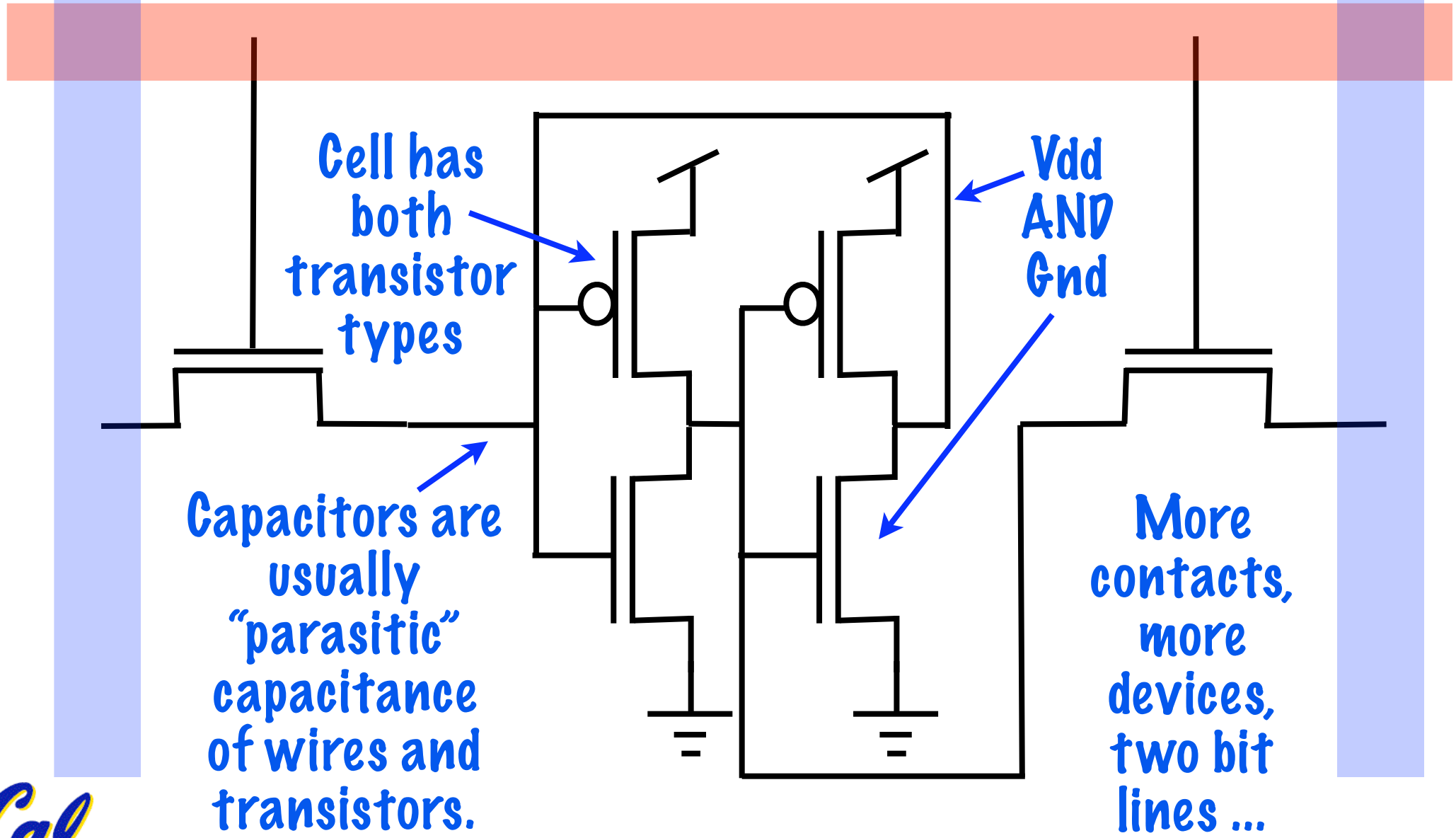
Combine both cases to complete circuit



UC Regents Fall 2005 © UCB

SRAM Challenge #1: It's so big!

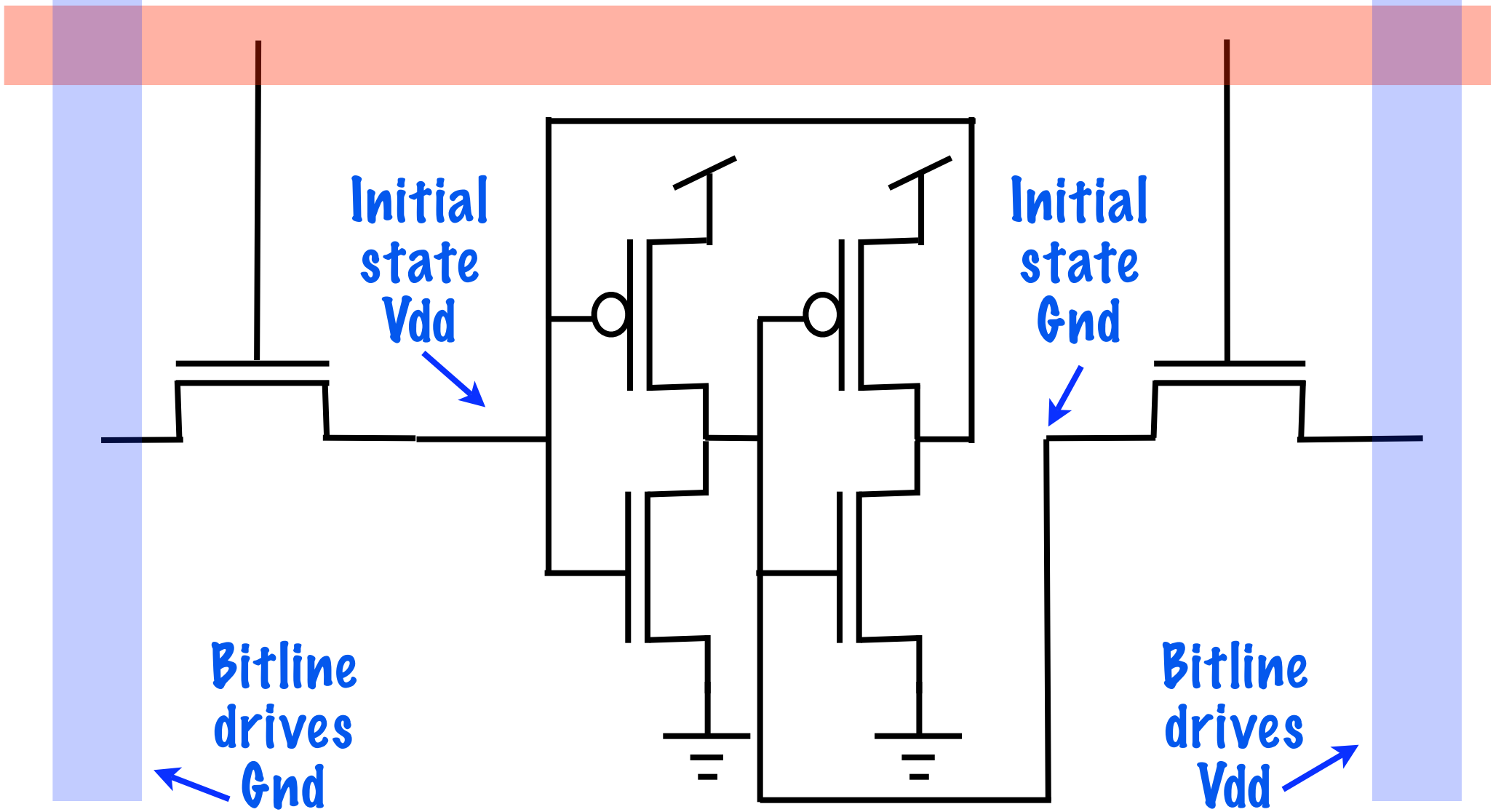
SRAM area is 6X-10X DRAM area, same generation ...



UC Regents Fall 2005 © UCB

Challenge #2: Writing is a "fight"

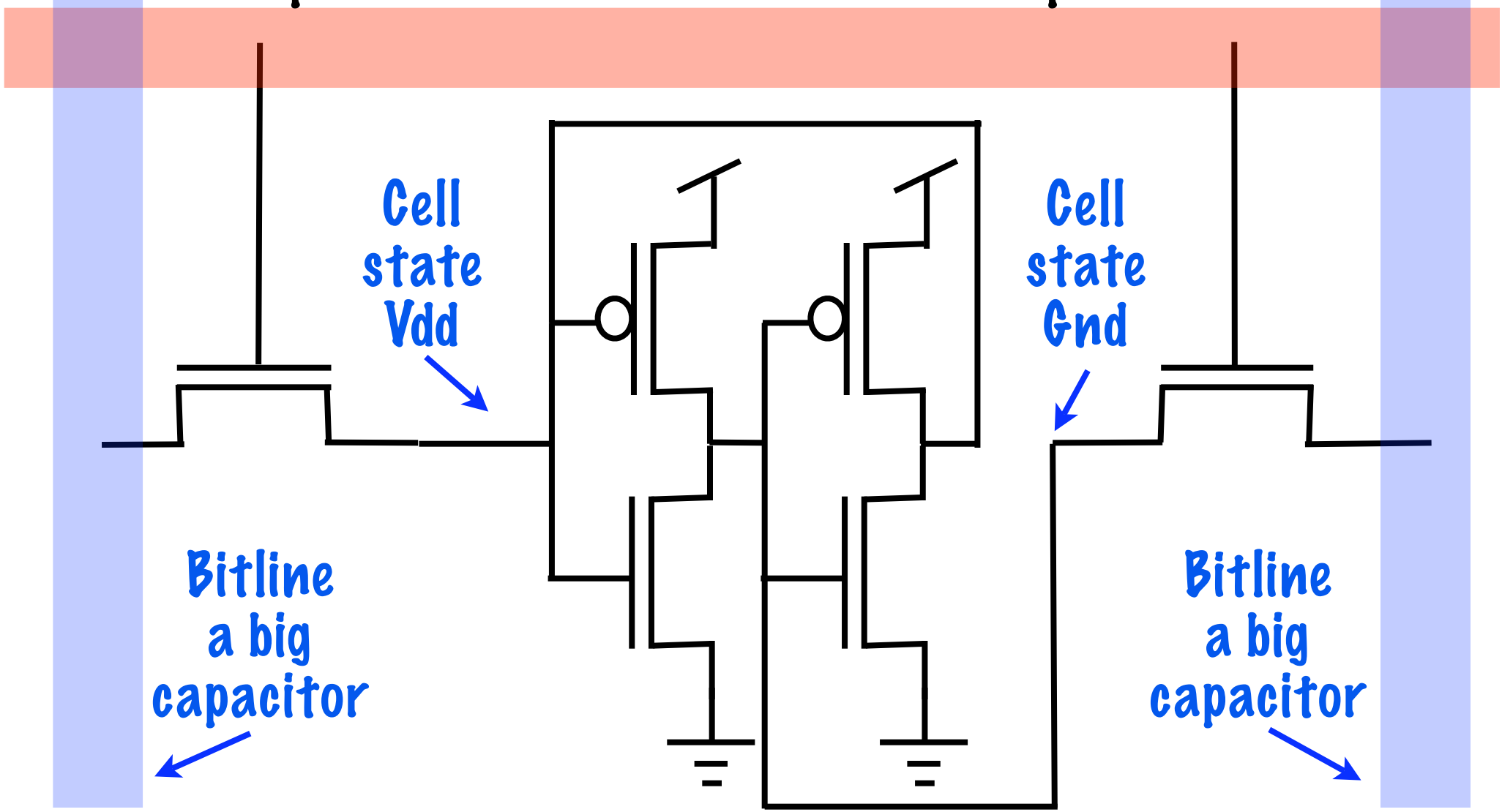
When word line goes high, bitlines "fight" with cell inverters to "flip the bit" -- must win quickly!
Solution: tune W/L of cell & driver transistors



UC Regents Fall 2005 © UCB

Challenge #3: Preserving state on read

When word line goes high on read, cell inverters must drive large bitline capacitance quickly, to preserve state on its small cell capacitances



UC Regents Fall 2005 © UCB

SRAM vs DRAM, pros and cons

Big win for DRAM

- **DRAM has a 6-10X density advantage at the same technology generation.**

SRAM advantages

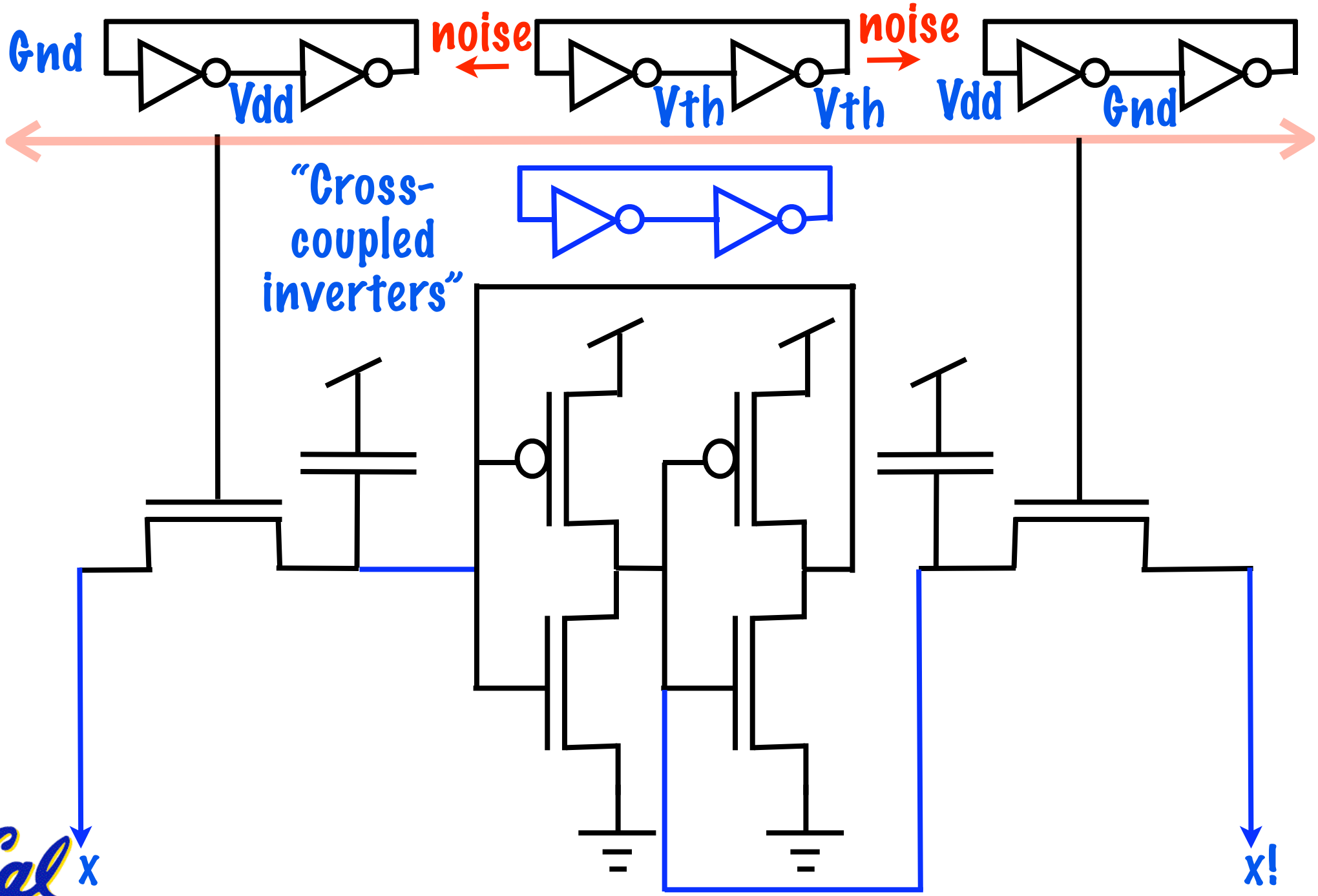
- **SRAM has deterministic latency: its cells do not need to be refreshed.**
- **SRAM is much faster: transistors drive bitlines on reads.**
- **SRAM easy to design in logic fabrication process (and premium logic processes have SRAM add-ons)**

UC Regents Fall 2005 © UCB

Flip Flops Revisited

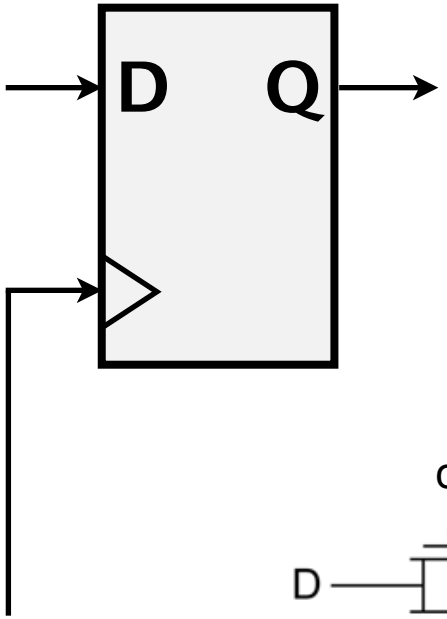
UC Regents Fall 2005 © UCB

Recall: Static RAM cell (6 Transistors)

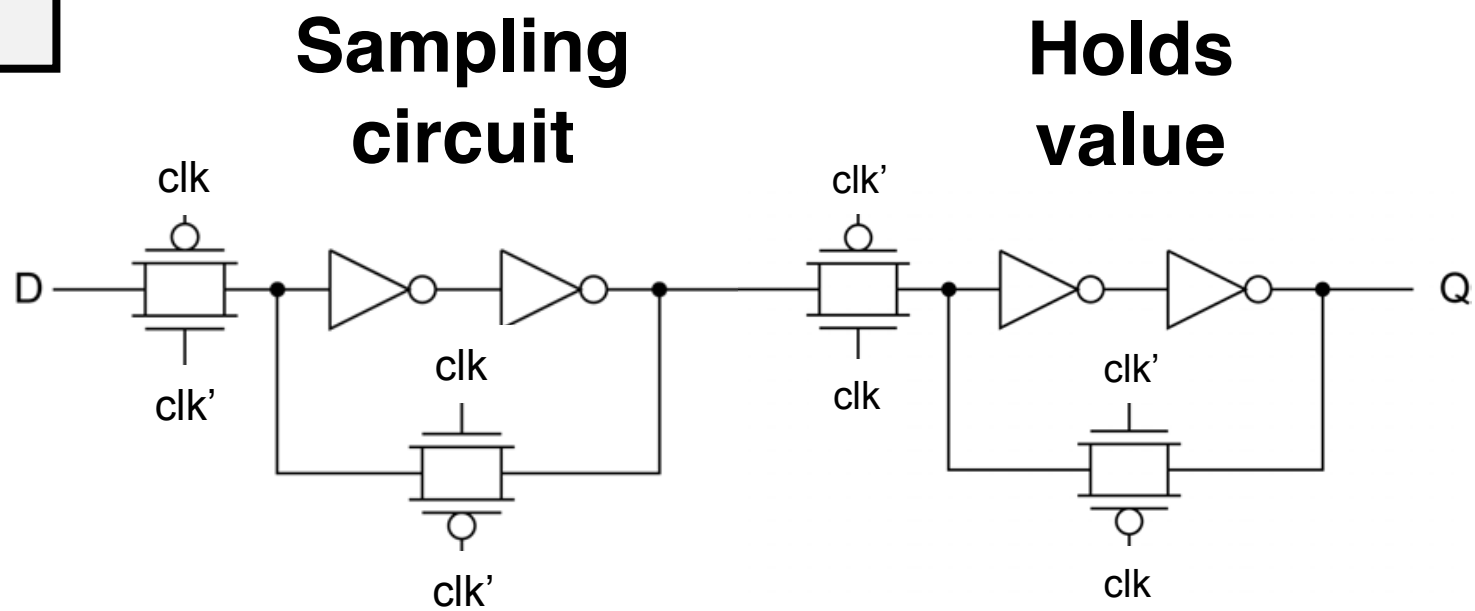


UC Regents Fall 2005 © UCB

Recall: Positive edge-triggered flip-flop



A flip-flop "samples" right before the edge, and then "holds" value.



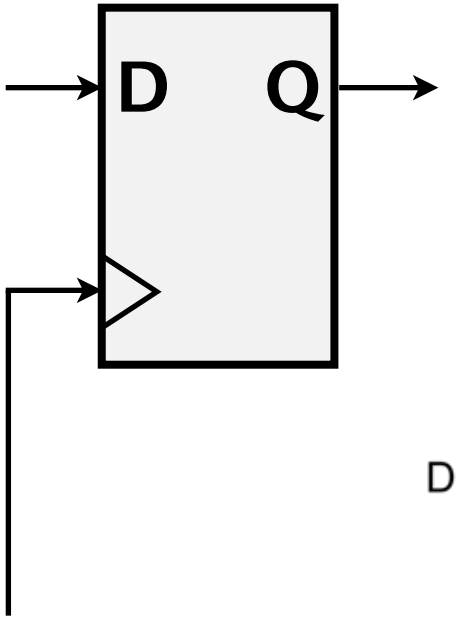
16 Transistors: Makes an SRAM look compact!

What do we get for the 10 extra transistors?

Clocked logic semantics.

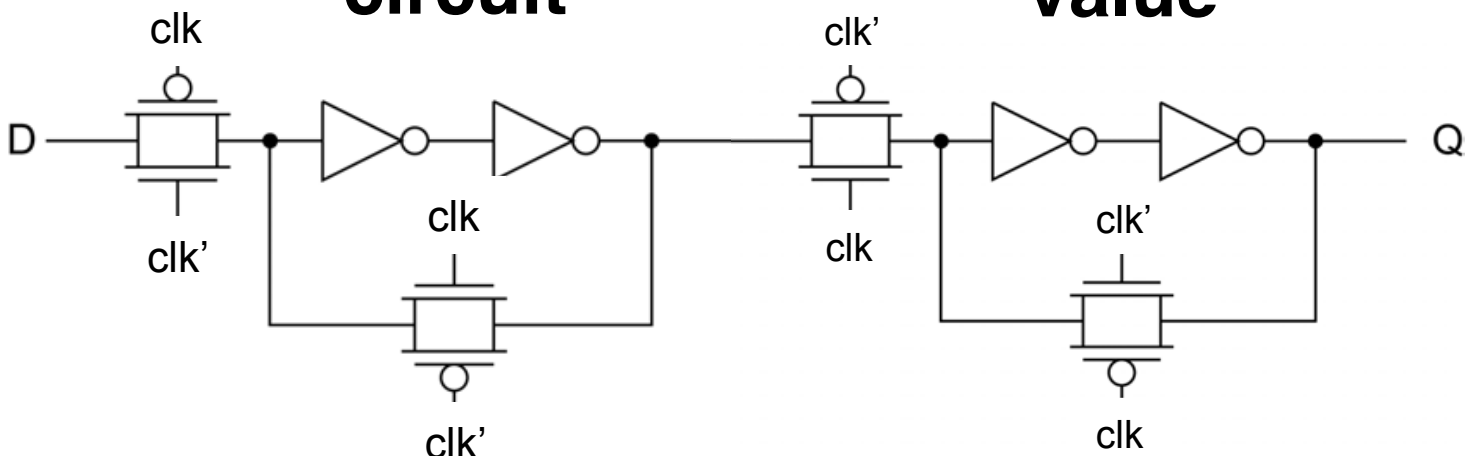
UC Regents Fall 2005 © UCB

Sensing: When clock is low

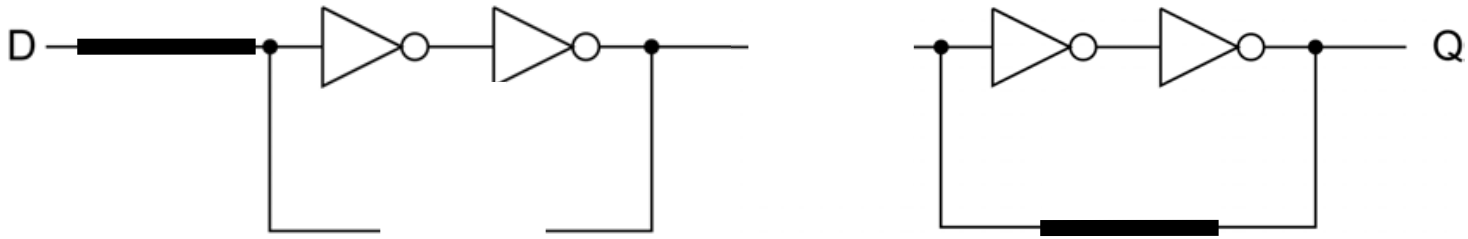


A flip-flop "samples" right before the edge, and then "holds" value.

Sampling circuit | **Holds value**



clk = 0
clk' = 1

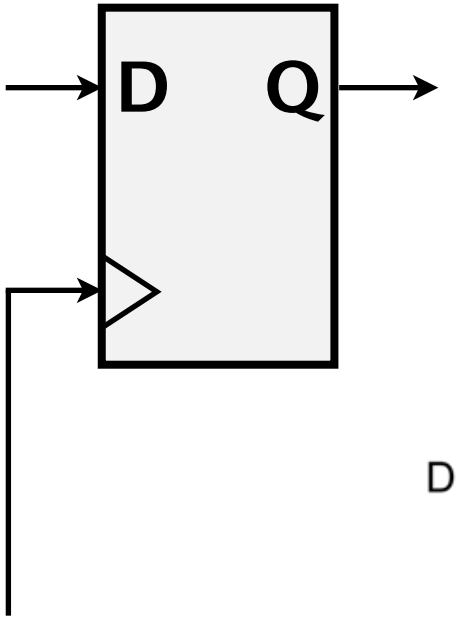


Will capture new value on posedge.

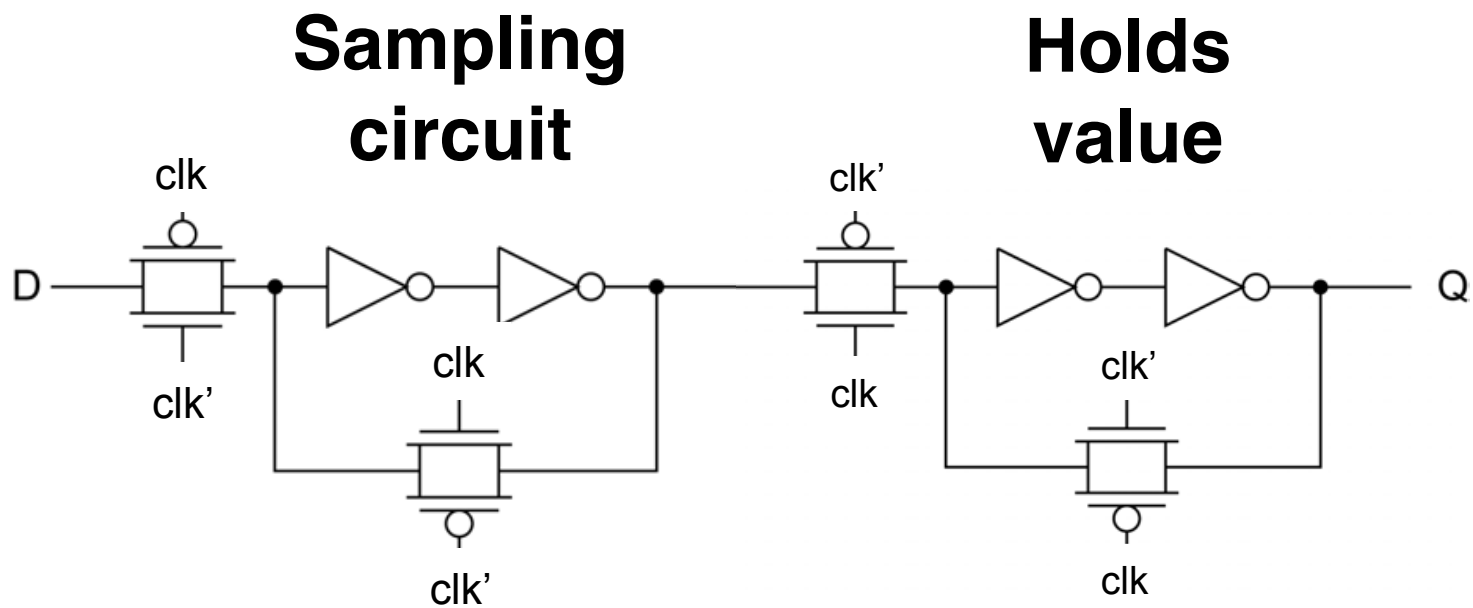
Outputs last value captured.

UC Regents Fall 2005 © UCB

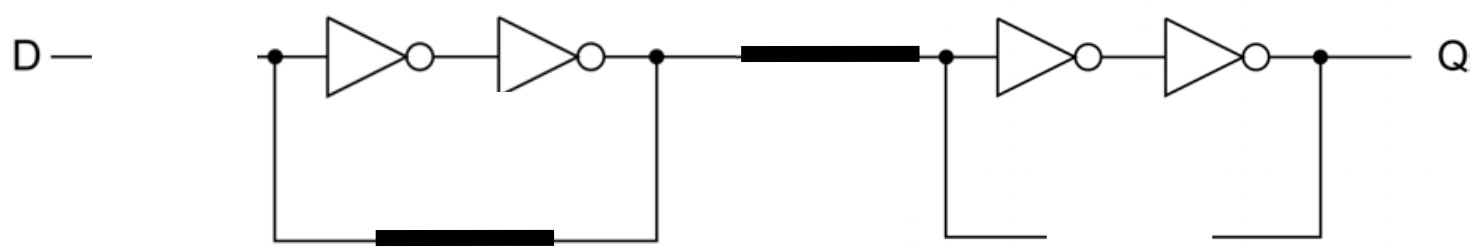
Capture: When clock goes high



A flip-flop "samples" right before the edge, and then "holds" value.



clk = 1
clk' = 0



Remembers value just captured.

Outputs value just captured.

UC Regents Fall 2005 © UCB

Admin: Final Xilinx Checkoff Friday ...

Th 10/13	Cache I	PDF	7.1-2	Preview Final Project
F 10/14				Lab 3: Final Xilinx Checkoff, 12-2PM or 3-5PM, 125 Cory

4. On **Friday 10/14** in lab section, you will demonstrate the processor running on the Calinx board. Unlike the 10/7 checkoff, this demo will test all facets of the processor (including the issues discussed in Problem 4). During this demo, the TA will provide you with secret test code. If you are able to pass these tests on your first try, you will receive bonus points. You can also receive bonus points if you fix your processor to pass the tests within your section time. If your processor is not fixed by the end of section, your TA will provide source for the secret test code, for use in your weekend debugging sessions.

M 10/17				Lab 3: Final Report Due, 11:59 PM via the submit program

Lab report due Monday, 11:59 PM.

UC Regents Fall 2005 © UCB

Memory Arrays

Calinx DRAM: 133 Mhz, 128 Mb

Micron™

128Mb: x4, x8, x16
SDRAM

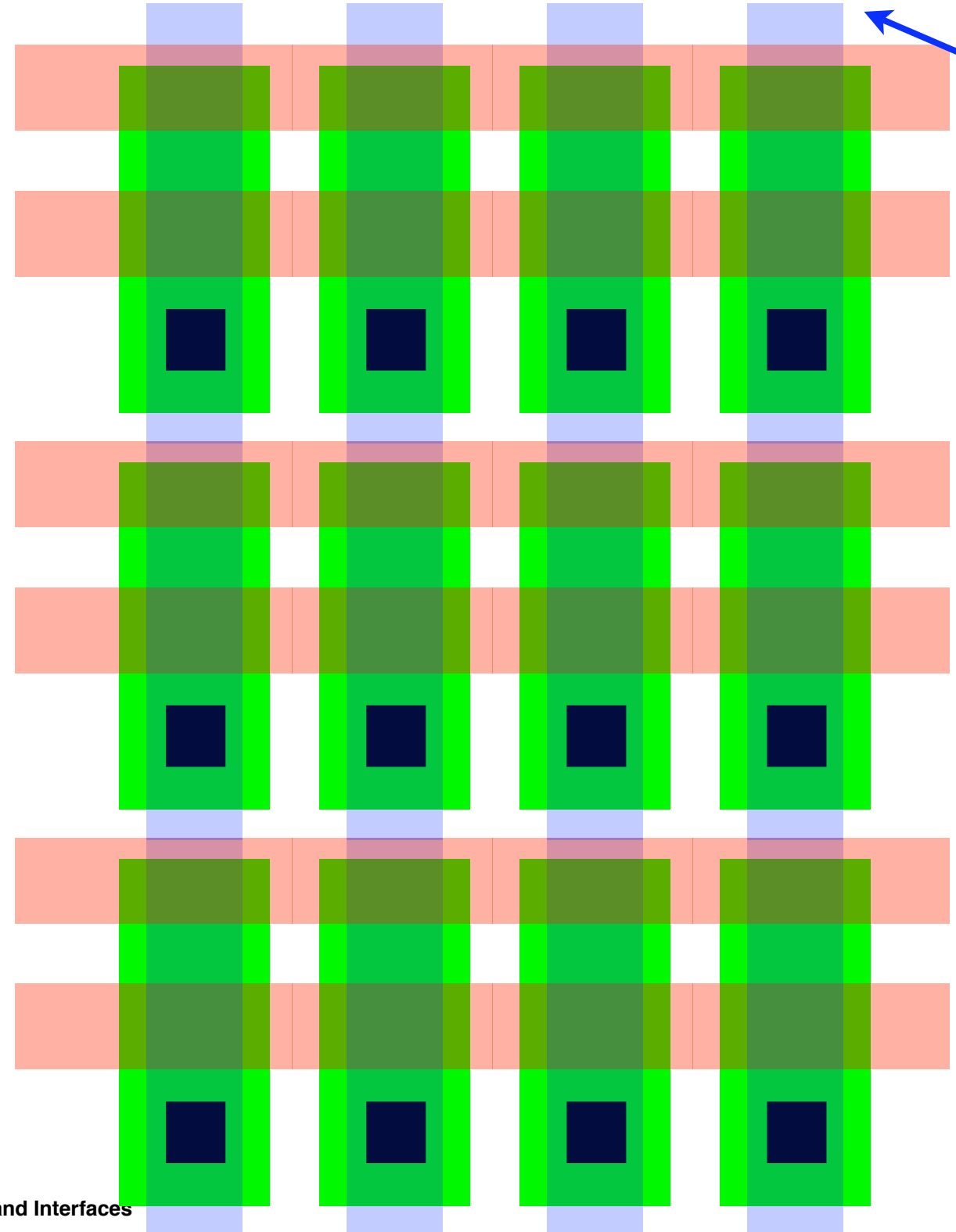
SYNCHRONOUS DRAM

MT48LC32M4A2 – 8 Meg x 4 x 4 banks
MT48LC16M8A2 – 4 Meg x 8 x 4 banks
MT48LC8M16A2 – 2 Meg x 16 x 4 banks

For the latest data sheet, please refer to the Micron Web site: www.micron.com/dramds

Data sheet on "resources" page. Will need to understand for final project!

UC Regents Fall 2005 © UCB

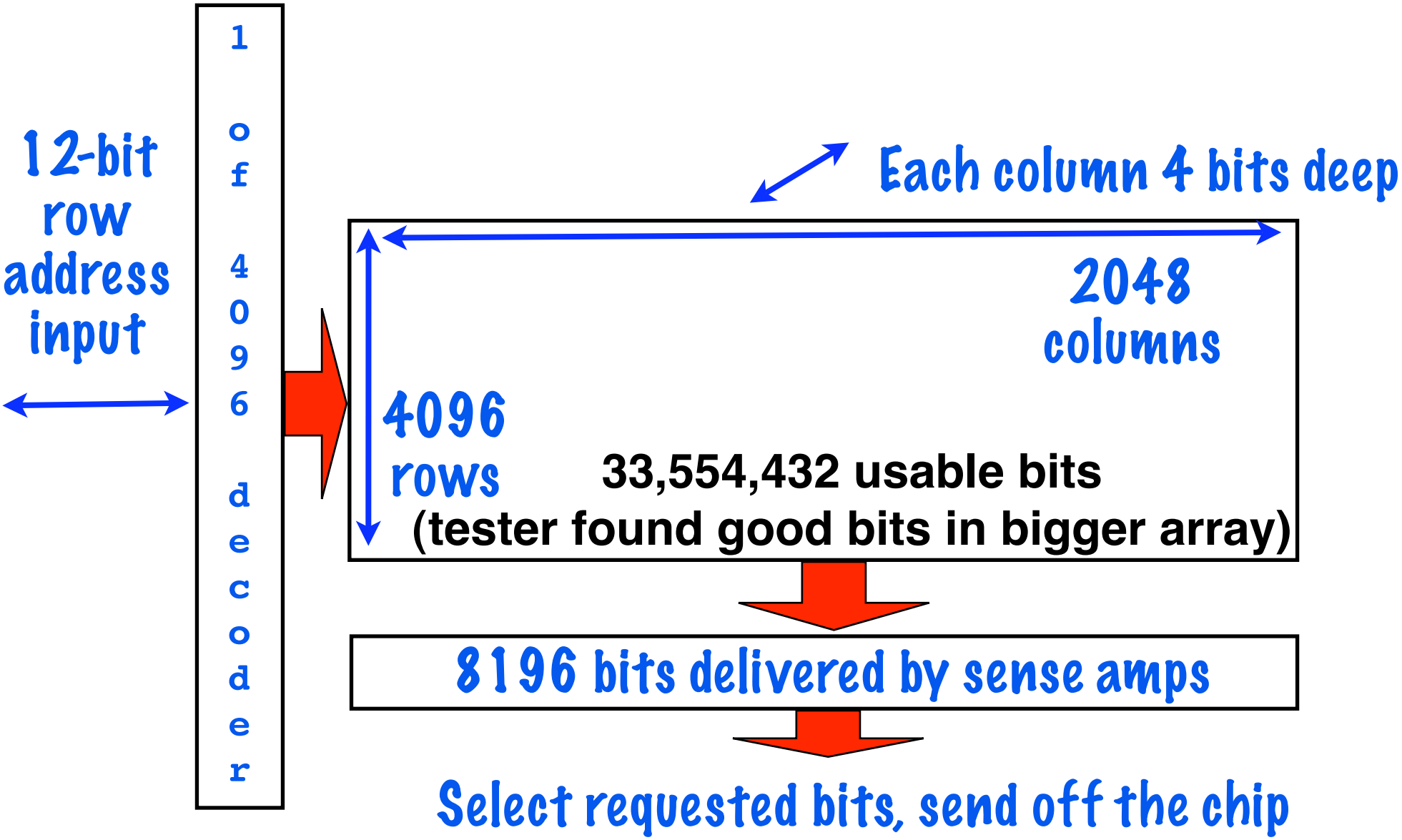


People buy DRAM for the bits. "Edge" circuits are overhead

So, we amortize the edge circuits over big arrays

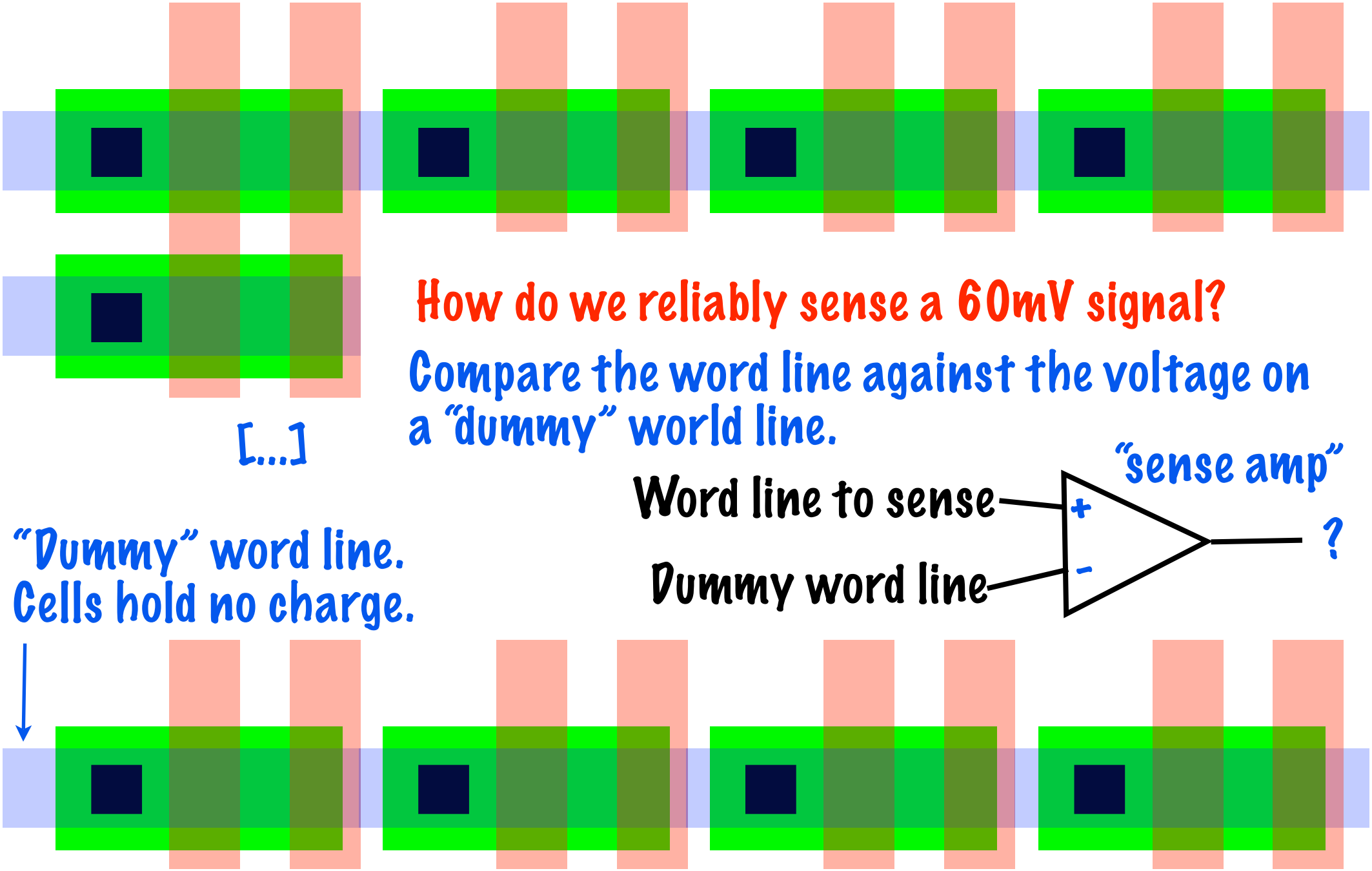
UC Regents Fall 2005 © UCB

A "bank" of 32 Mb (128Mb chip -> 4 banks)



UC Regents Fall 2005 © UCB

Recall DRAM Challenge #3b: Sensing

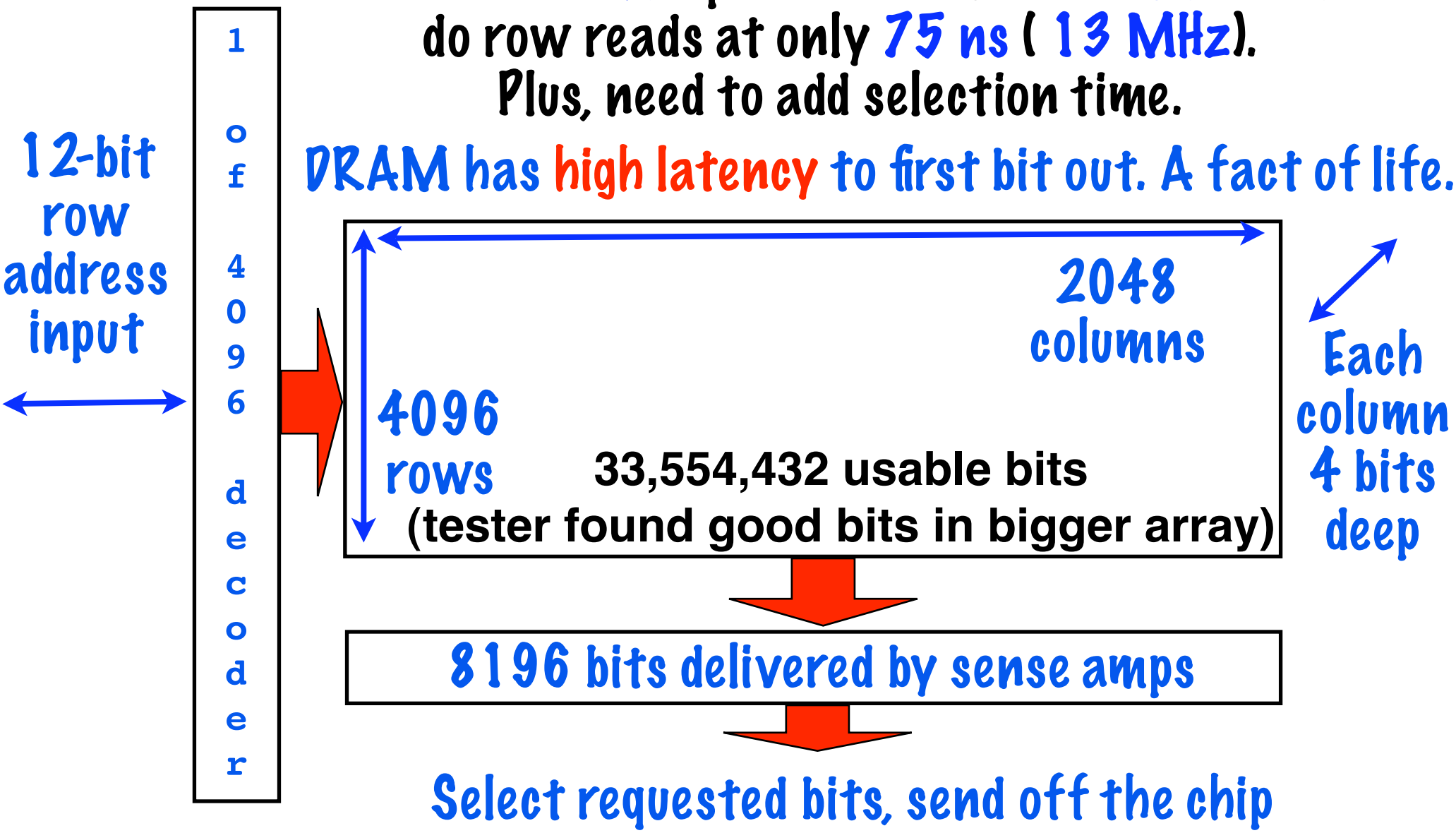


UC Regents Fall 2005 © UCB

Corresponds to row read into sense amps

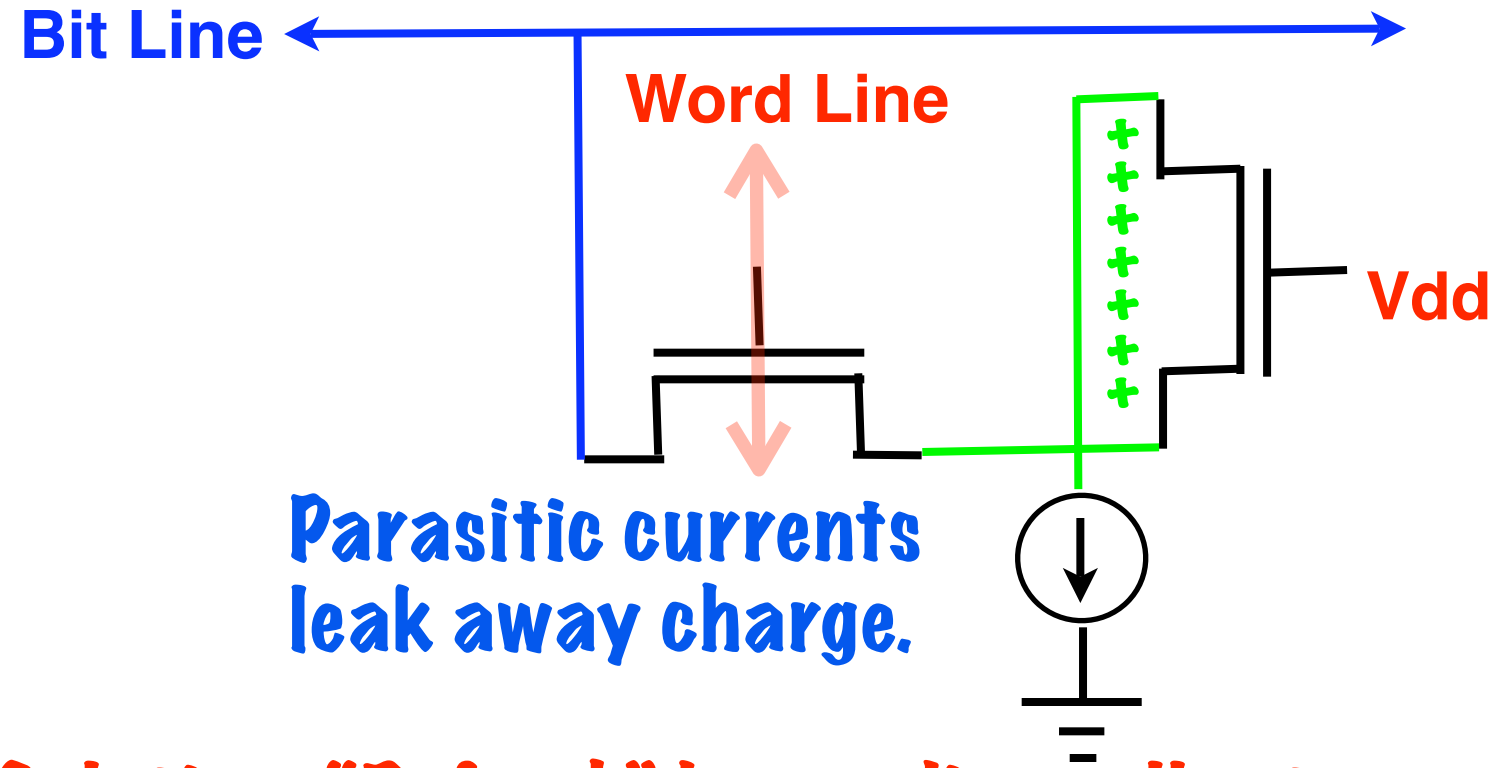
Slow! This 7.5ns period DRAM (133 MHz) can do row reads at only 75 ns (13 MHz). Plus, need to add selection time.

DRAM has high latency to first bit out. A fact of life.

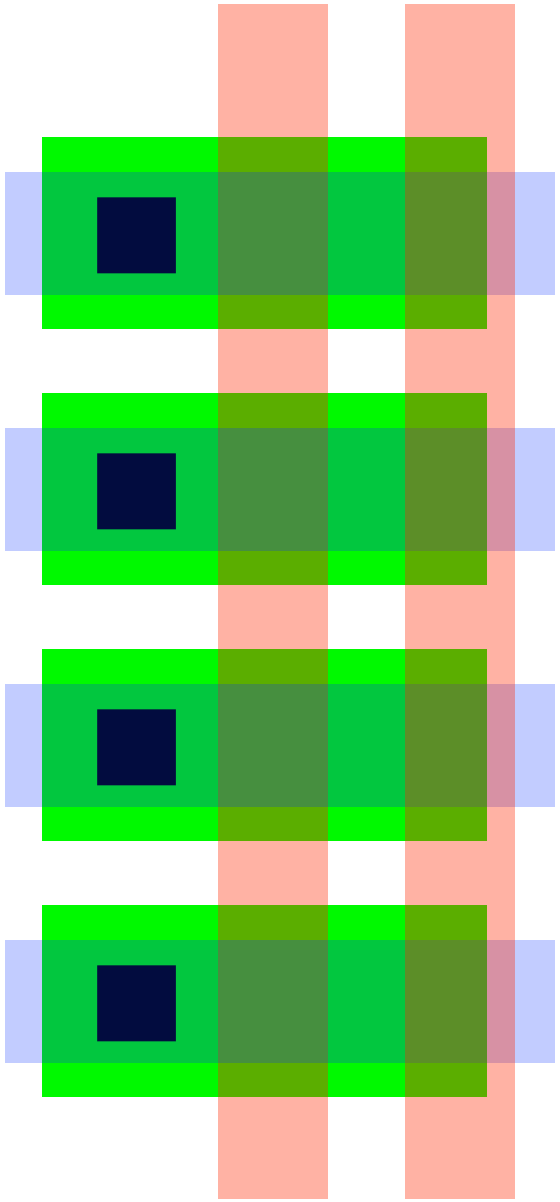


UC Regents Fall 2005 © UCB

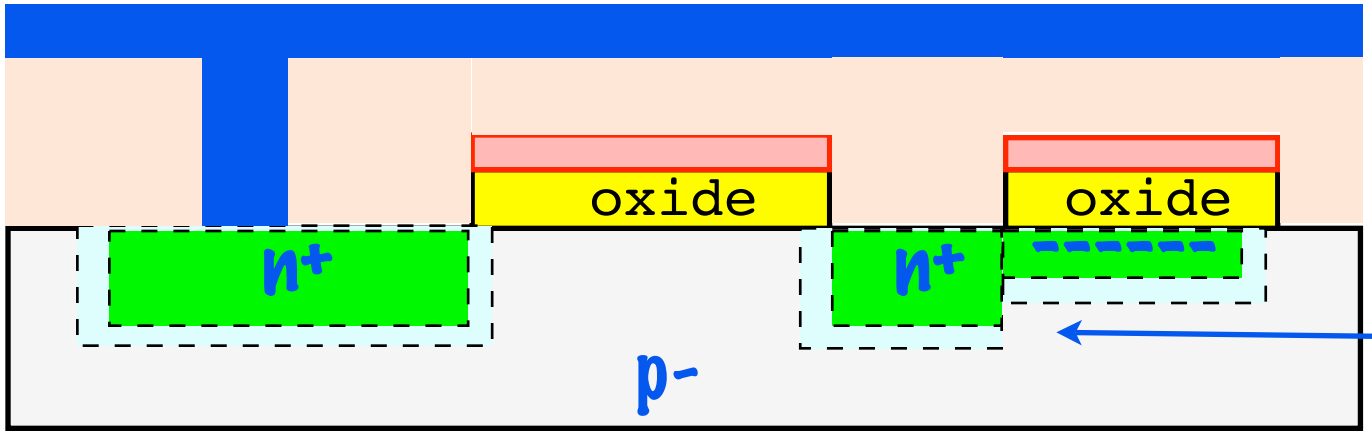
An ill-timed refresh may add to latency



Parasitic currents leak away charge.



Solution: "Refresh", by reading cells at regular intervals (tens of milliseconds)



Diode leakage ...

UC Regents Fall 2005 © UCB

Latency is not the same as bandwidth!

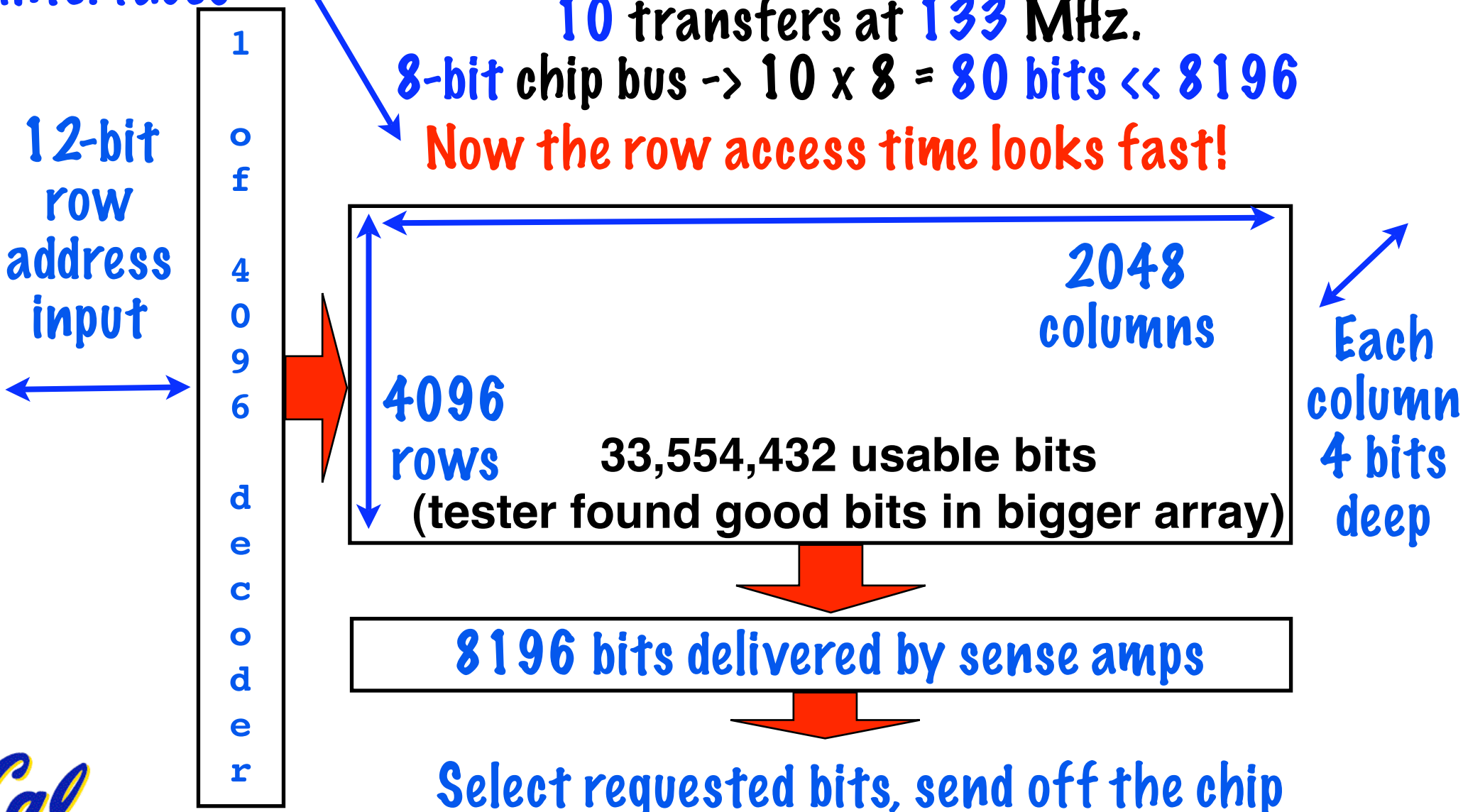
Thus, push to faster DRAM interfaces

What if we want all of the 8196 bits?

In row access time (75 ns) we can do 10 transfers at 133 MHz.

8-bit chip bus -> 10 x 8 = 80 bits << 8196

Now the row access time looks fast!



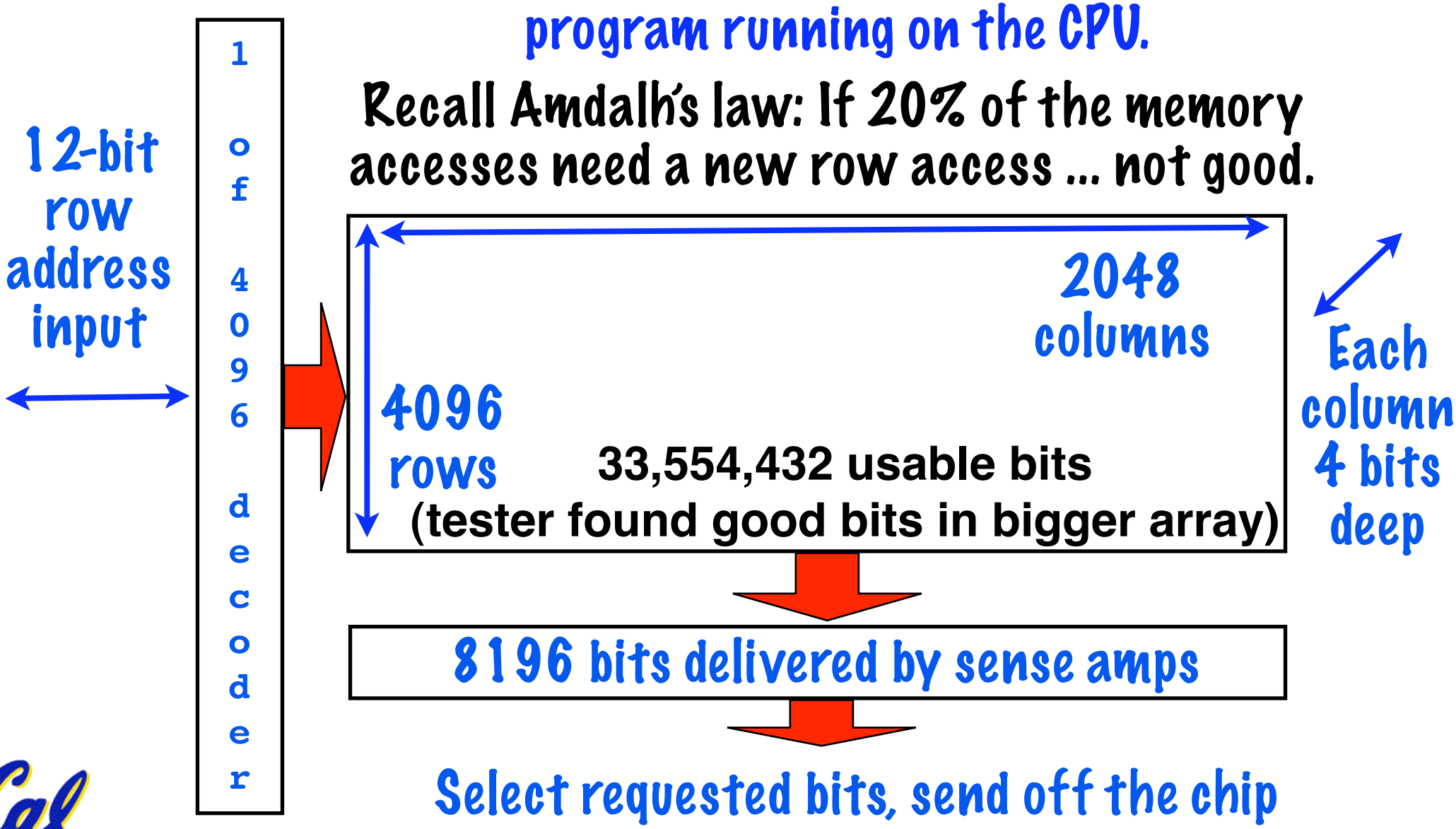
UC Regents Fall 2005 © UCB

Sadly, it's rarely this good ...

What if we want all of the 8196 bits?

The "we" for a CPU would be the program running on the CPU.

Recall Amdalh's law: If 20% of the memory accesses need a new row access ... not good.



UC Regents Fall 2005 © UCB

DRAM latency/bandwidth chip features

- **Columns:** Design the right interface for CPUs to request the subset of a column of data it wishes:

8196 bits delivered by sense amps

Select requested bits, send off the chip

- **Interleaving:** Design the right interface to the 4 memory banks on the chip, so several row requests run in parallel.

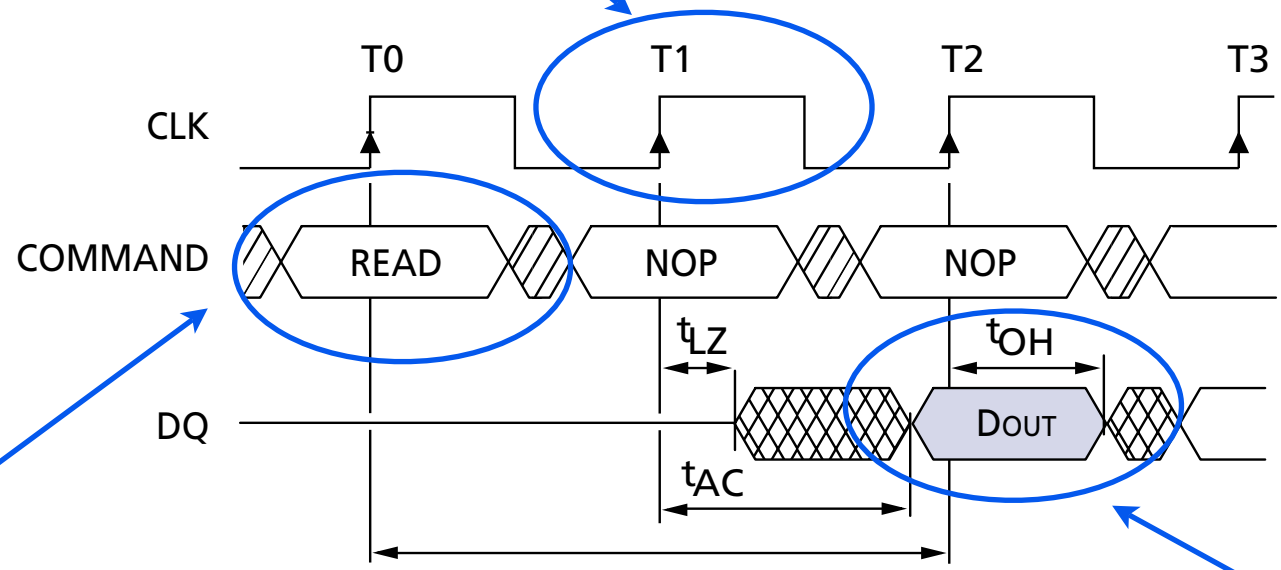
Bank 1 | Bank 2 | Bank 3 | Bank 4

UC Regents Fall 2005 © UCB

Off-chip interface for the Micron part ...

A clocked bus protocol (133 MHz)

Note! This example is best-case! For a random access, DRAM takes many more than 2 cycles!



CAS Latency = 2
(CAS = Column Address Strobe)

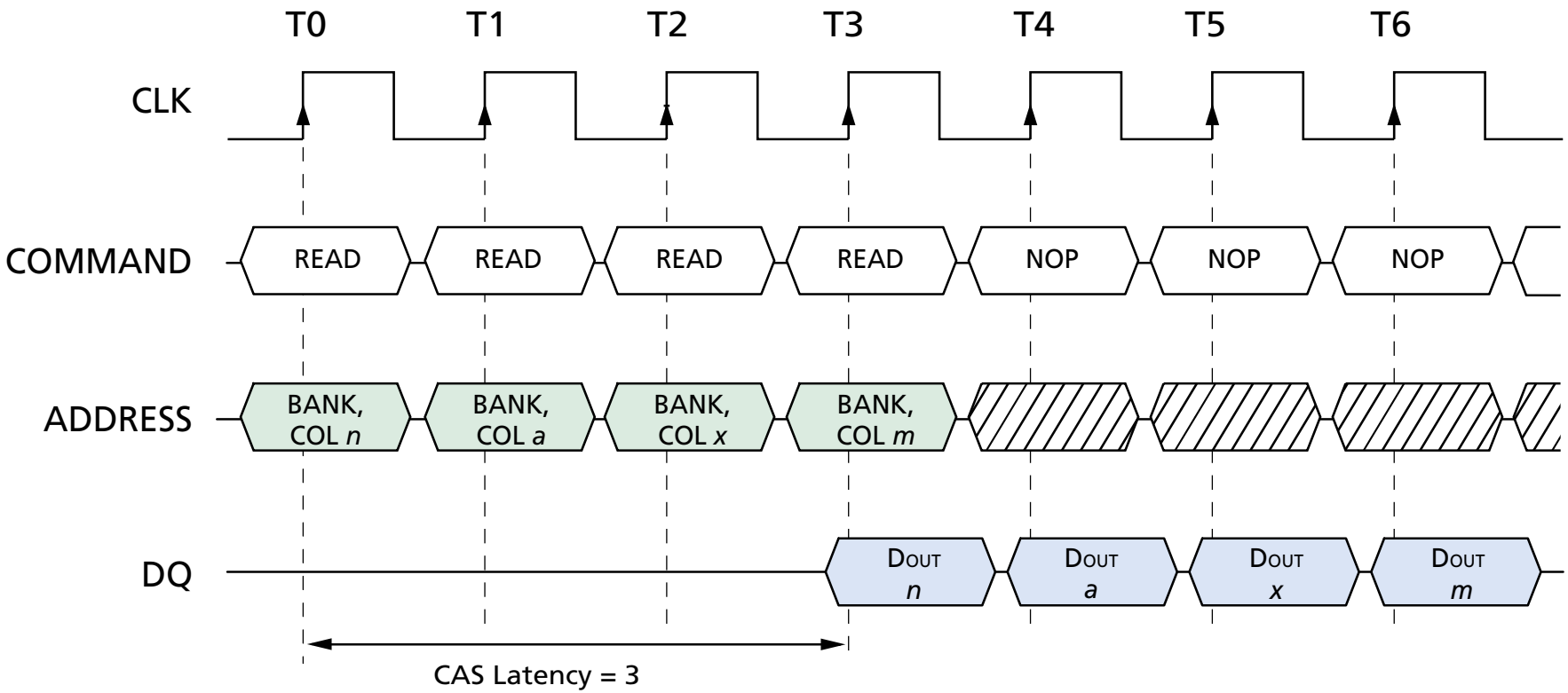
DRAM is controlled via commands (READ, WRITE, REFRESH, ...)

Synchronous data output with variable latency

From Micron 128 Mb SDRAM data sheet (on "resources" web page)

UC Regents Fall 2005 © UCB

Example: Access all 4 banks in parallel



NOTE: Each READ command may be to any bank. DQM is LOW.

DON'T CARE

Figure 8
Random READ Accesses

UC Regents Fall 2005 © UCB

Lectures: Coming up next ...

Th 10/13	Cache I
F 10/14	
Sa 10/15	
Su 10/16	
M 10/17	
T 10/18	Cache II

Essential tools for the final project.

UC Regents Fall 2005 © UCB