United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1977-8

First ICTP Regional Microelectronics Workshop and Training on VHDL for Hardware Synthesis and FPGA Design in Asia-Pacific

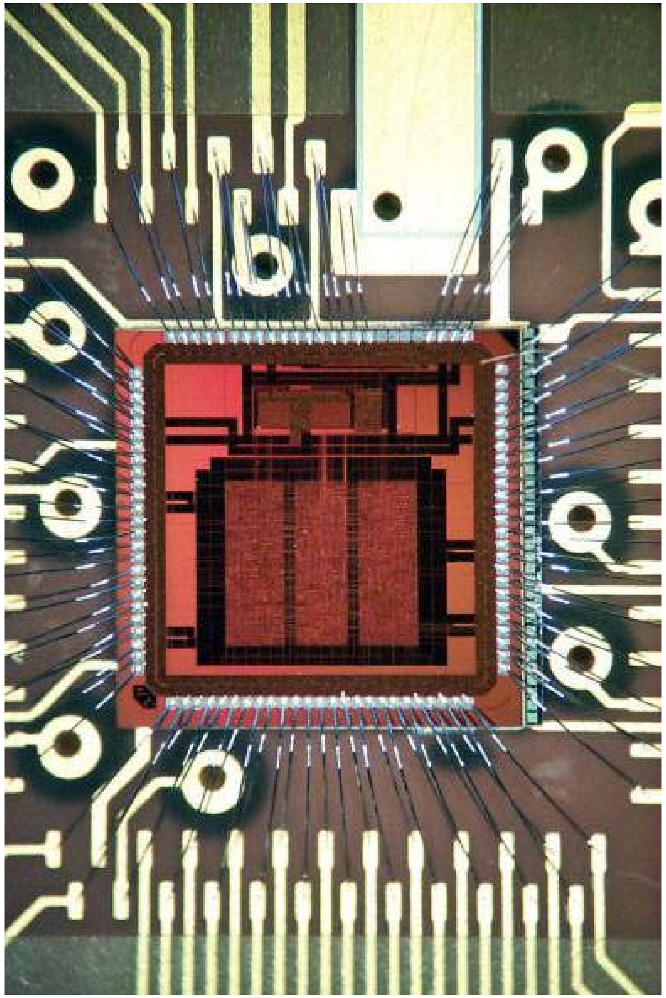
16 June - 11 July, 2008

Storage elements.

Paulo Moreira
PH ESE ME Division
C E R N
CH-1211 Geneva 23
SWITZERLAND

Outline

- Introduction
- Transistors
- The CMOS inverter
- Technology
- Scaling
- Gates
- Sequential circuits
- Storage elements
  - Read-only
  - Nonvolatile R/W
  - Read-write
    - 6T SRAM
    - 3T dynamic
    - 1T dynamic
- Phase-Locked Loops
- Example

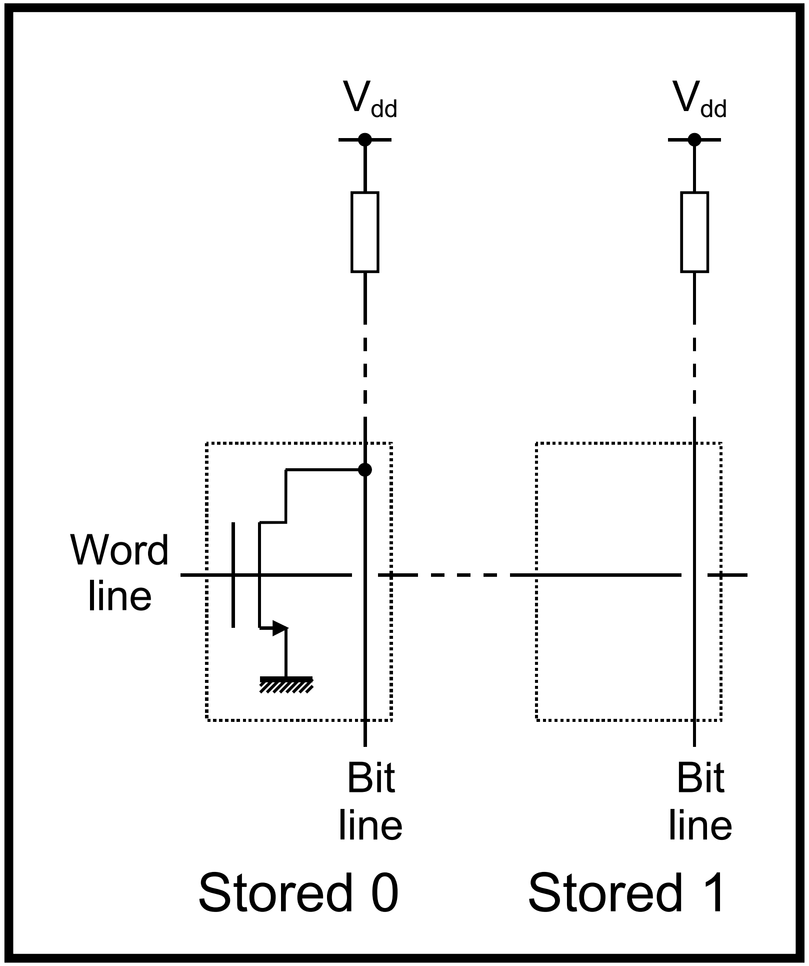


Storage elements

- The silicon area of large memory cells is dominated by the size of the memory core, it is thus crucial to keep the size of the basic storage cell as small as possible
- The storage cell area is reduced by:
  - reducing the driving capability of the cell (small devices)
  - reducing the logic swing and the noise margins
- Consequently, sense amplifiers are used to restore full rail-to-rail amplitude

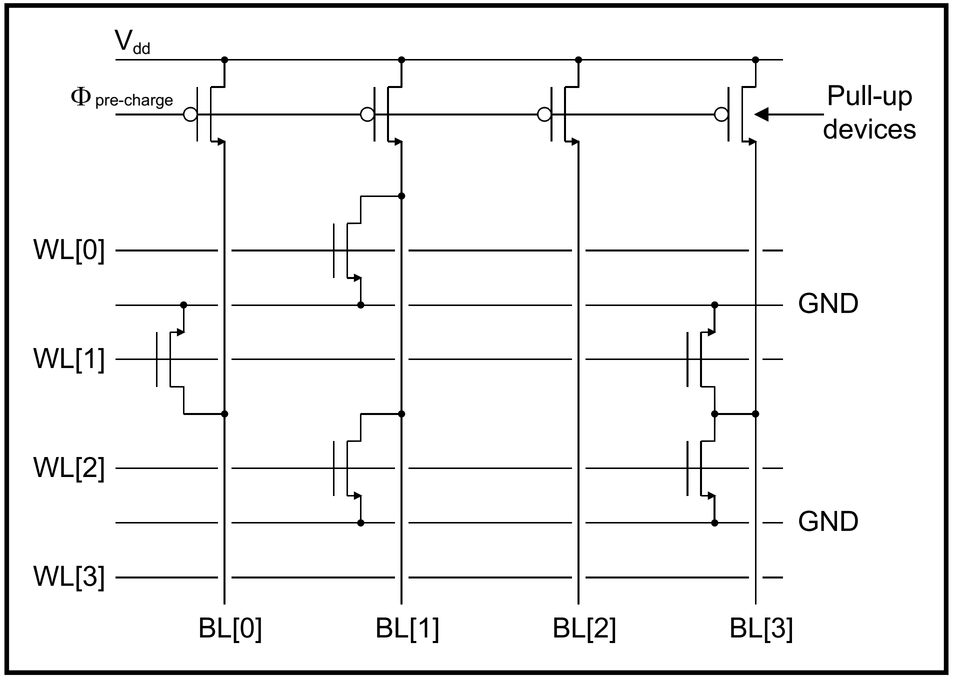
Read-only

- Because the contents is permanently fixed the cell design is simplified
- Upon activation of the word line a 0 or 1 is presented to the bit line:
  - If the NMOS is absent the word line has no influence on the bit line:
    - The word line is pulled-up by the resistor
    - A 1 is stored in the “cell
  - If the NMOS is present the word line activates the NMOS:
    - The word line is pulled-down by the NMOS
    - A 0 is stored in the cell



Read-only

- In practice a "always on" pull-up device is never used because:
  - V_{OL} would depend on the ratio of the pull-up/pull-down devices
  - A static current path would exists when the output is low causing high power dissipation in large memories
- In practice pre-charged logic is used:
  - Eliminates the static dissipation
  - Pull-up devices can be made wider
- The bit lines are first pre-charged by the pull-up devices
  - during this phase the word lines must be disabled
- Then, the word lines are activated (word evaluation)
  - during this phase the pull-up devices are off



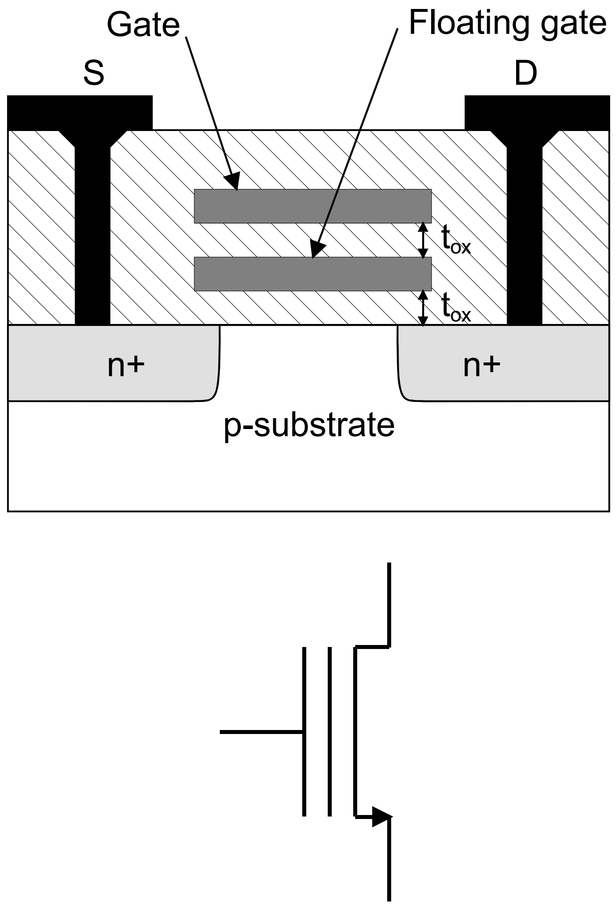
Nonvolatile R/W

- The same architecture as a ROM memory
- The pull-down device is modified to allow control of the threshold voltage
- The modified threshold is retained "indefinitely":
  - The memory is nonvolatile
- To reprogram the memory the programmed values must be erased first
- The "heart" of NVRW memories is the Floating Gate Transistor (FAMOS)

Nonvolatile R/W

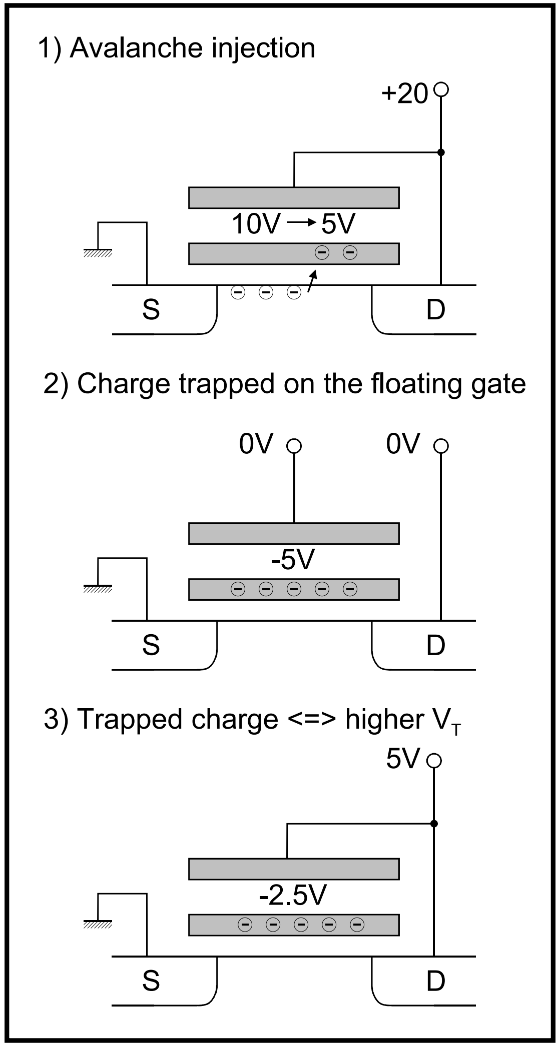
- A floating gate is inserted between the gate and the channel
- The device acts as a normal transistor
- However, its threshold voltage is programmable
- Since the t_{ox} is doubled, the transconductance is reduced to half and the threshold voltage increased

Floating-gate transistor (FAMOS)



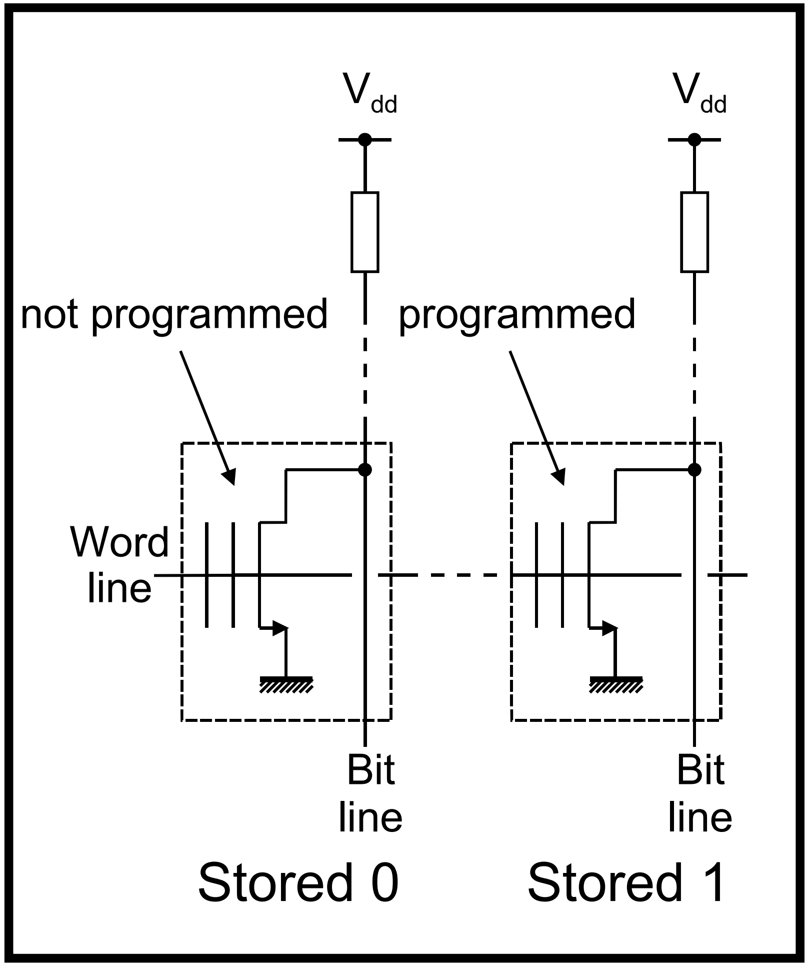
Nonvolatile R/W

- Programming the FAMOS:
  - A high voltage is applied between the source and the gate-drain
  - A high field is created that causes avalanche injection to occur
  - Electrons traverse the first oxide and get trapped on the floating gate (t_{ox} = 100nm)
  - Trapped electrons effectively drop the floating gate voltage
  - The process is self limiting: the building up of gate charge eventually stops avalanche injection
  - The FAMOS with a charged gate is equivalent to a higher V_T device
  - Normal circuit voltages can not turn a programmed device on



Nonvolatile R/W

- The non-programmed device can be turned on by the word line thus, it stores a "0"
- The word line high voltage can not turn on the programmed device thus, it stores a "1"
- Since the floating gate is surrounded by SiO_2, the charge can be stored for many years

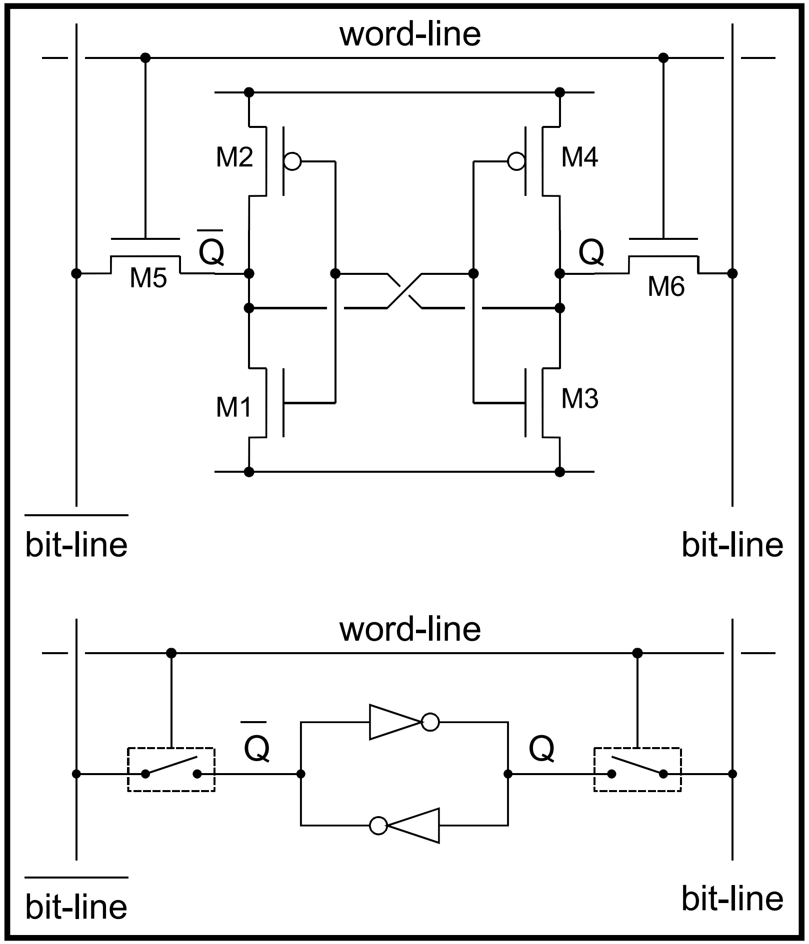


Nonvolatile R/W

- Erasing the memory contents (EPROM):
  - Strong UV light is used to erase the memory:
    - UV light renders the oxide slightly conductive by direct generation of electron-hole pairs in the SiO_2
  - The erasure process is slow (several minutes)
  - Programming takes 5-10μs/word
  - Number of erase/program cycles limited (<1000)
- Electrically-Erasable PROM (E^2PROM)
  - A reversible tunneling mechanism allows E^2PROM's to be both electrically programmed and erased

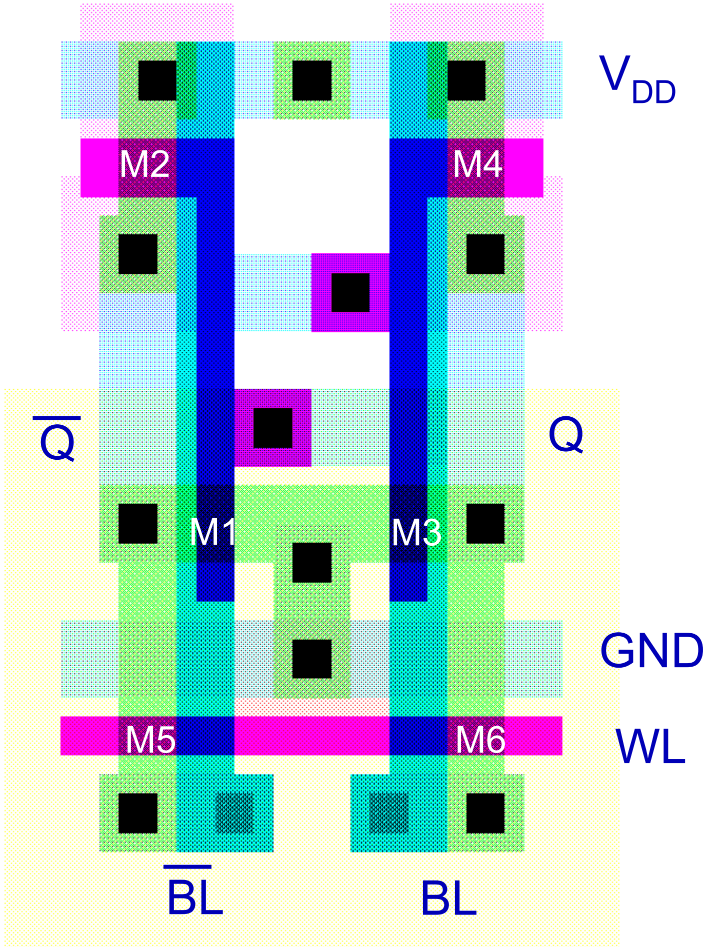
6T SRAM

- Static Read-Write Memories (SRAM):
  - data is stored by positive feedback
  - the memory is volatile
- The cell use six transistors
- Read/write access is enabled by the word-line
- Two bit lines are used to improve the noise margin during the read/write operation
- During read the bit-lines are pre-charged to $V_{dd}/2$:
  - to speedup the read operation
  - to avoid erroneous toggling of the cell



6T SRAM

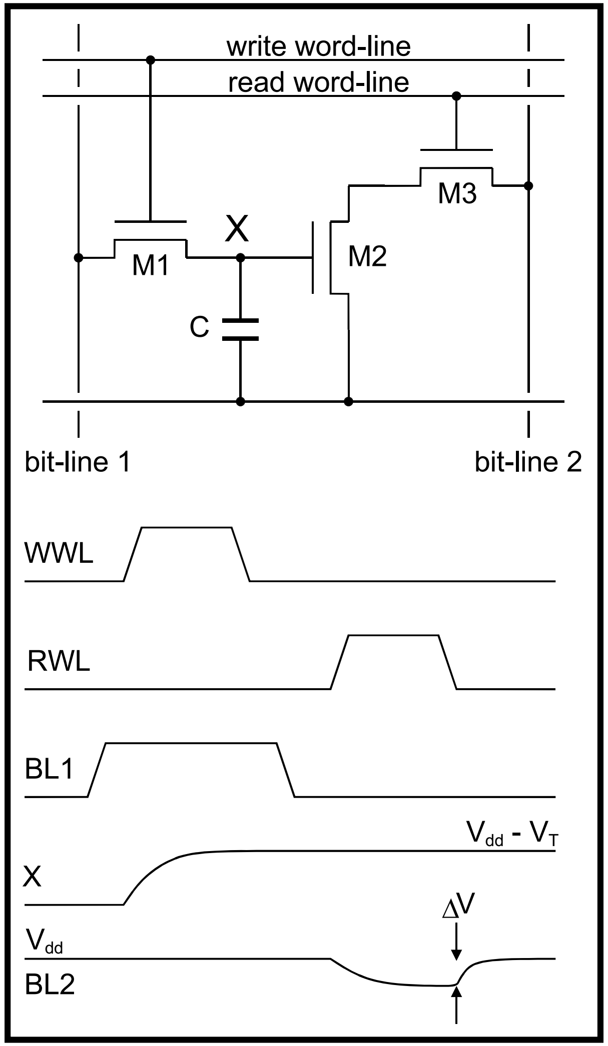
- SRAM performance:
  - The read operation is the critical one:
    - It involves discharging or charging the large bit-line capacitance through the small transistors of the cell
  - The write time is dominated by the propagation delay of the cross-coupled inverter pair
  - The six-transistor cell is not area efficient:
    - It requires routing of two power lines, two bit lines and a word line
    - Most of the area is taken by wiring and interlayer contacts



(from J. M. Rabaey 1996)

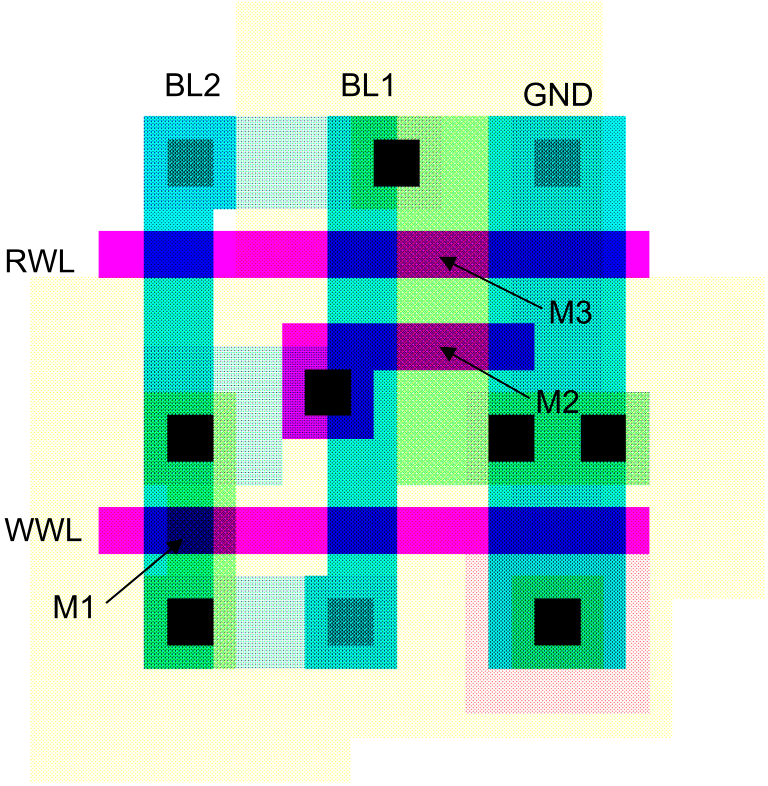
3T Dynamic

- Dynamic Random-Access Memory (DRAM)
  - In a dynamic memory the data is stored as charge in a capacitor
- Tree-Transistor Cell (3T DRAM):
  - Write operation:
    - Set the data value in bit-line 1
    - Assert the write word-line
    - Once the WWL is lowered the data is stored as charge in C
  - Read operation:
    - The bit-line BL2 is pre-charged to Vdd
    - Assert the read word-line
    - if a 1 is stored in C, M2 and M3 pull the bit-line 2 low
    - if a 0 is stored C, the bit-line 2 is left unchanged



3T Dynamic

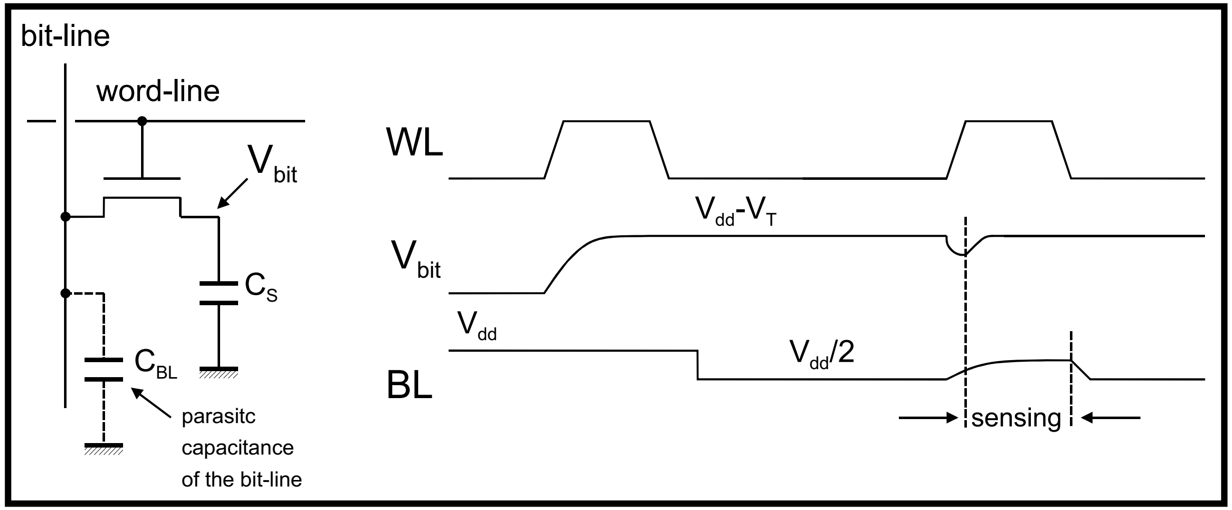
- The cell is inverting
- Due to leakage currents the cell needs to be periodically refreshed (every 1 to 4ms)
- Refresh operation:
  - read the stored data
  - put its complement in BL1
  - enable/disable the WWL
- Compared with an SRAM the area is greatly reduce:
  - SRAM ⇒ 1092 λ^2
  - DRAM ⇒ 576 λ^2
  - The area reduction is mainly due to the reduction of the number of devices and interlayer contacts



(from J. M. Rabaey 1996)

1T Dynamic

- One-Transistor dynamic cell (1T DRAM)
  - It uses a single transistor and a capacitor
  - It is the most widely used topology in commercial DRAM's
- Write operation:
  - Data is placed on the bit-line
  - The word-line is asserted
  - Depending on the data value the capacitance is charged or discharged

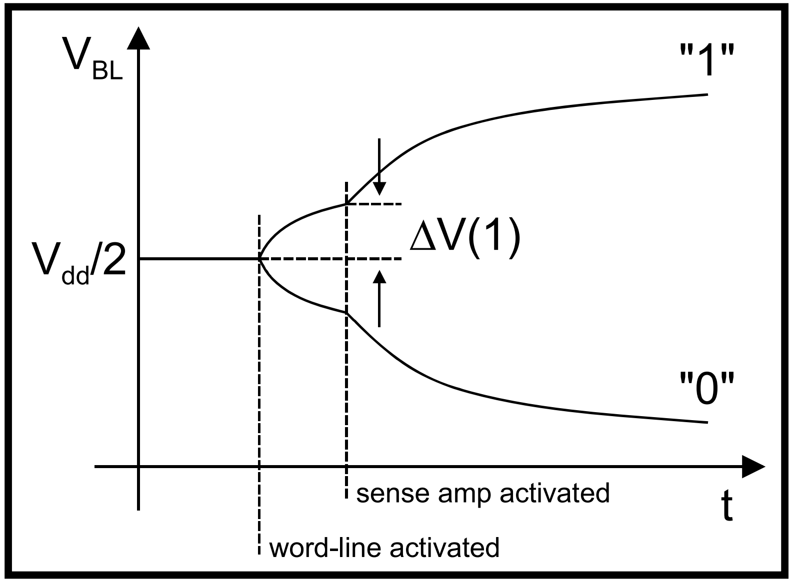


1T Dynamic

- Read operation:
  - The bit-line is pre-charged to $V_{dd}/2$
  - The word-line is activated and charge redistribution takes place between C_S and the bit-line
  - This gives origin to a voltage change in the bit-line, the sign of which determines the data stored:

$$\Delta V = \left(V_{BIT} - \frac{V_{dd}}{2}\right)\frac{C_S}{C_S + C_{BL}}$$

  - C_{BL} is 10 to 100 times bigger than $C_S \Rightarrow \Delta V \cong 250mV$

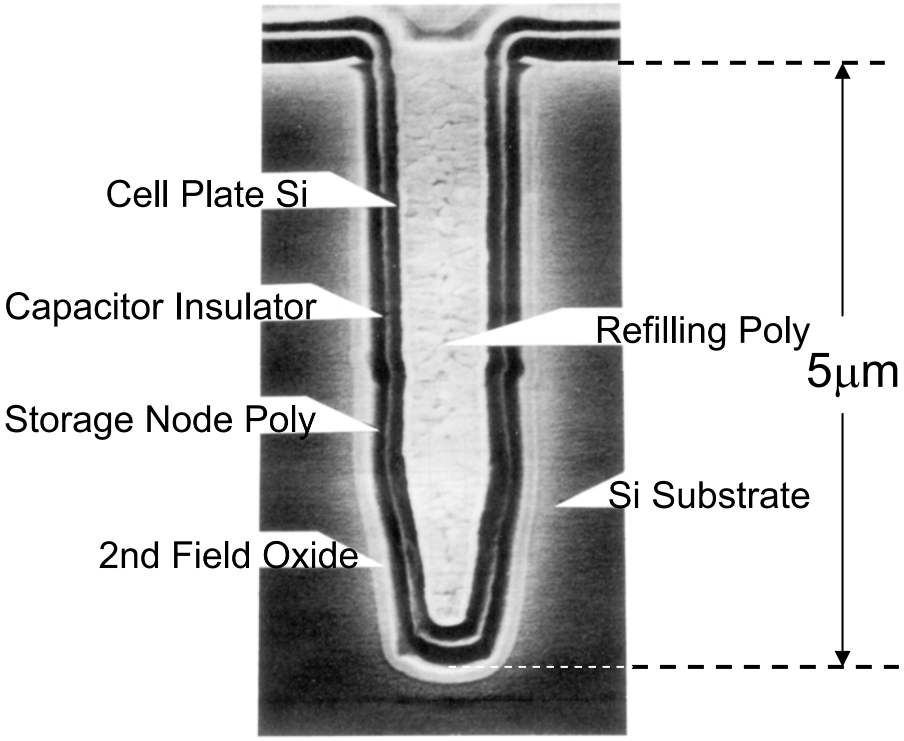


  - The amount of charge stored in the cell is modified during the read operation
  - However, during read, the output of the sense amplifier is imposed on the bit line restoring the stored charge

1T Dynamic

- Contrary to the previous cases a 1T cell requires a sense amplifier for correct operation
- Also, a relatively large storage capacitance is necessary for reliable operation
- A 1 is stored as $V_{dd}-V_T$. This reduces the available charge:
  - To avoid this problem the word-line can be bootstrapped to a value higher than Vdd

Trench capacitor



(from T. Mano et al., 1987)